

Result No.	Score	Query Match	Length	DB	ID	Description
1	741.5	37.4	399	4	US-09-491-562-7	Sequence 7, Appl
2	741.5	37.4	399	4	US-09-874-562-7	Sequence 7, Appl
3	741.5	37.4	477	4	US-09-449-335-6	Sequence 6, Appl
4	737.5	37.2	477	4	US-09-449-335-6	Sequence 2, Appl
5	737.5	37.2	477	4	US-09-480-921B-29	Sequence 29, Appl
6	736	37.1	477	4	US-09-480-921B-10	Sequence 10, Appl
7	654.5	33.0	475	4	US-09-491-562-2	Sequence 2, Appl
8	654.5	33.0	475	4	US-09-874-562-2	Sequence 2, Appl
9	138.5	7.0	2293	4	US-09-368-590-2	Sequence 2, Appl
10	115	5.8	5215	4	US-09-105-537-2	Sequence 2, Appl
11	113.5	5.7	7257	3	US-09-335-409-5	Sequence 5, Appl
12	113.5	5.7	7257	4	US-09-568-102-5	Sequence 5, Appl
13	113.5	5.7	7257	4	US-09-567-969-5	Sequence 5, Appl
14	113.5	5.7	7257	4	US-09-568-480-5	Sequence 5, Appl
15	113.5	5.7	7257	4	US-09-568-486-5	Sequence 5, Appl
16	113.5	5.7	7257	4	US-09-568-472-5	Sequence 5, Appl
17	113.5	5.7	7257	4	US-09-567-899-5	Sequence 5, Appl
18	111	5.6	5087	4	US-09-144-085-1	Sequence 1, Appl
19	110.5	5.6	3729	2	US-08-804-227C-4	Sequence 1, Appl
20	109.5	5.5	4472	2	US-08-804-227C-2	Sequence 2, Appl
21	108.5	5.5	3724	2	US-08-804-227C-10	Sequence 10, Appl
22	108.5	5.5	3724	2	US-08-804-198-4	Sequence 4, Appl
23	107.5	5.4	6095	4	US-09-144-085-2	Sequence 2, Appl
24	106.5	5.4	609	4	US-08-969-683A-67	Sequence 67, Appl
25	103.5	5.2	4551	3	US-09-320-878-1	Sequence 1, Appl
26	103.5	5.2	4613	4	US-09-105-537-6	Sequence 31, Appl
27	103.5	5.2	11877	4	US-09-105-537-6	Sequence 6, Appl

28	102	5.1	1346	3	US-09-320-878-4	Sequence 4, Appl1
29	102	5.1	1346	4	US-09-105-537-37	Sequence 3, Appl1
30	102	5.1	1580	2	US-08-804-222C-11	Sequence 11, Appl1
31	102	5.1	1580	2	US-08-804-199-5	Sequence 5, Appl1
32	99.5	5.0	3170	4	US-09-036-987A-4	Sequence 4, Appl1
33	99.5	5.0	3170	4	US-09-370-700-4	Sequence 4, Appl1
34	98.5	5.0	419	1	US-08-305-172B-4	Sequence 4, Appl1
35	98	4.9	1382	3	US-09-057-570-4	Sequence 4, Appl1
36	97.5	4.9	3739	3	US-09-320-878-2	Sequence 2, Appl1
37	97.5	4.9	3739	3	US-09-105-537-33	Sequence 3, Appl1
38	97	4.9	3170	2	US-07-642-734C-5	Sequence 5, Appl1
39	97	4.9	3170	3	US-08-443-009A-5	Sequence 5, Appl1
40	96.5	4.9	523	2	US-08-923-772-2	Sequence 2, Appl1
41	96.5	4.9	523	4	US-09-385-287-2	Sequence 2, Appl1
42	96.5	4.9	2152	4	US-09-036-987A-3	Sequence 3, Appl1
43	96.5	4.9	2152	4	US-09-370-700-3	Sequence 3, Appl1
44	95.5	4.8	1864	2	US-08-804-222C-4	Sequence 3, Appl1
45	95.5	4.8	3519	4	US-09-428-517-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
 US-09-491-362-7
 Sequence 7, Application US/09491362
 Patent No. 6281012
 GENERAL INFORMATION:
 APPLICANT: Croteau, Rodney B
 APPLICANT: Lange, Bernd M
 TITLE OF INVENTION: 1-DEOXY-D-XYULOSE-5-PHOSPHATE REDUCTOISOMERASE, AND
 TITLE OF INVENTION: METHODS OF USE
 FILE REFERENCE: MSUR14977
 CURRENT APPLICATION NUMBER: US/09/491,362
 CURRENT FILING DATE: 2000-01-26
 EARLIER APPLICATION NUMBER: 60/118,349
 EARLIER FILING DATE: 1999-02-03
 NUMBER OF SEQ ID NOS: 13
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 7
 LENGTH: 399
 TYPE: PRT
 ORGANISM: Arabidopsis thaliana
 US-09-491-362-7

Query Match	37.48;	Score 741.5;	DB 4;	Length 399;
-------------	--------	--------------	-------	-------------

Matches 162; Conservative 71; Mismatches 140; Indels 19; Gaps 6;

Qy	2	RSLSITGAGSGISESFDUMMKGGEPFVALATGCRNIRBLAEARALKLWETANE	61
		: : : : : : : : : : : : : : : : : : : : : : : : :	
Dd	3	KPISIVSGTSGTQYLDIVAE - NPDKFRVVALAAGSNVTTLLADQYRRFKPALVAARNE	60
Qy	62	DCPLREALAGTG--TEVAGCAQAIAMADRP-ADMTSIAITGAAGLVGPMALKHGT	111
		: : : : : : : : : : : : : : : : : : : : : : : : :	
Dd	61	SLINEKELEADLDYKLEITPEQDQVIEVABHPRAVTVVVGIGCAGLRYAALAEAGD	120
Qy	119	LALANKESIVTAGOLIMRAOENCATILPVBSEHSAFQALAGEPRACERYIITASGP	178
		: : : : : : : : : : : : : : : : : : : : : : : : :	
Dd	121	IALANKETLLAGRPFLPLANKHNKYLIPADSEHSALFQCIOGLPGCALRKIITLASGA	181
Qy	179	FRWMSLEIRIACVAAEAQHPRMSMGQRISIDSMSMKLDELITREFGEPRIEAV	233
		: : : : : : : : : : : : : : : : : : : : : : : : :	
Dd	181	FRMPVEKLEKVVADALKNHPNMNKGKTIYDSATFLNKLEVEIENHYLFGAEYDIEIV	241
Qy	239	VHDSIVLHAAVGFCDGJLANHIGRAPDMKHAIGFALMPGR---GEYPAVARIDALIASLT	291
		: : : : : : : : : : : : : : : : : : : : : : : : :	
Dd	241	IHPQSIHSMIETGROSSVLAQGMWDMRLPILYTMSMPDRVPDSEVTWRDLCKTGLST	300
Qy	296	FQRPDEERRPALRLARDVMAARGLSGAFAFNAAKTALDHTIAGRGFIDMAAIVEETL--	355
		: : : : : : : : : : : : : : : : : : : : : : : : :	
Dd	301	EKKRPDVKKPSMDLIVAAAGRAGGTVGLTGLSAAANKAAVEKIDKISYLDIFVVEVLTCDK	360

QY 354 --AGVSTDPLEFGKVPDALEEVLAAMDHLARRAA 383
Db 361 HRNELVTSP-----SLEEIVHYDLMAREYA 385

RESULT 2

US-09-874-562-7

Sequence 7, Application US/09874562

Patent No. 6420159

GENERAL INFORMATION:

APPLICANT: Croteau, Rodney B

APPLICANT: Lange, Bernd M

TITLE OF INVENTION: 1-DEOXY-D-XYLULOSE-5-PHOSPHATE REDUCTOISOMERASE, AND

TITLE OF INVENTION: METHODS OF USE

FILE REFERENCE: WSUR17549

CURRENT APPLICATION NUMBER: US/09/874,562

CURRENT FILING DATE: 2001-06-04

PRIOR APPLICATION NUMBER: 09/491,362

PRIOR FILING DATE: 2000-01-26

PRIOR APPLICATION NUMBER: 60/118,349

PRIOR FILING DATE: 1999-02-03

NUMBER OF SEQ ID NOS: 13

SOFTWARE: Patentln Ver. 2.0

SEQ ID NO 7

LENGTH: 399

TYPE: PRT

ORGANISM: Arabidopsis thaliana

US-09-874-562-7

Query Match

Best Local Similarity 41.3%; Pred. No. 1.5e-67;

Matches 162; Conservative 71; Mismatches 140; Indels 19; Gaps 6;

QY 2 RSLIFGATGSGISTGPDVLRKGGPEAFRTVALTGRNIRRLAEMARALKAEIAYTAHE 61
Db 3 KRISIVGSGTIGTQTDIVAE--NPDKFRVVALAAGSNVTLADQVRRFPALVAYRNE 60

QY 62 DCLPALREALAGTG--TEVAGGAQALAEADRP--ADWTMSAIVGAAGLVPGMRALKHGR 118
Db 61 SLINELKEALADLDYKLEIIPGEGVIEVARHPRPAVTVTGICAGAKPTVAALIEAGKD 120

QY 119 LALANKESLVTAAGLMTAENGATILPVDESHSAVQALAGEDTCAVEYVITTAASGP 178
Db 121 IALANKETLNGPVPVPLANKHNKXILIPADSEHSALFOCIQGLEPAGALKRIITLASGA 180

QY 179 FRDMSLERIRACTVAEAOAHPMWSMGORISIDSASMEFKALELJETREFFGFEEDRIEAV 238
Db 181 FRDMPVERLKEKVKYADALKHNNMNGKKITVDSATLFNKGLEIVEAHLFGAEYDDIEIV 240

QY 239 VHPOSIYHAMVGFCDGGLMAHLGPADMRHAIGFALNMPGR--GEVYARIDLAQIASLT 295
Db 241 IHPOSIHSMIETODSSVLAQLGMPDMRLPILYTMSPDRVPCSEVYWPRLDCKLSLT 300

QY 296 FOKPDEERFPLRLAROVMAARGISGAFAFNKAEIALDHFIAGRIAGFLDMAVAVEEYL-- 353
Db 301 FKKPDNKKYPSMDLAIYAAGRGAGTGTGVLASANKAKAEMFIDEKISYIDIFKVELTCDK 360

QY 354 --AGVSTDPLEFGKVPDALEEVLAAMDHLARRAA 383
Db 361 HRNELVTSP-----SLEEIVHYDLMAREYA 385

RESULT 3
US-09-449-335-6
Sequence 6, Application US/09449335
Patent No. 6303365
GENERAL INFORMATION:
APPLICANT: Busch, Marco
APPLICANT: Hain, Rudiger
APPLICANT: Martin, William
APPLICANT: Tietjen, Klaus
APPLICANT: Kloti, Andreas
TITLE OF INVENTION: Method of determining the activity of

TITLE OF INVENTION: 1-deoxy-D-xylose-5-phosphate reductoisomerase and
TITLE OF INVENTION: 1-deoxy-D-xylose-5-phosphate synthase

FILE REFERENCE: 2020US

CURRENT APPLICATION NUMBER: US/09/449,335

CURRENT FILING DATE: 1999-11-24

EARLIER APPLICATION NUMBER: DE 199 35 967.9

EARLIER FILING DATE: 1999-07-30

NUMBER OF SEQ ID NOS: 8

SOFTWARE: Patentln Ver. 2.1

SEQ ID NO 6

LENGTH: 477

TYPE: PRT

ORGANISM: Arabidopsis thaliana

US-09-449-335-6

Query Match 37.4%; Score 741.5; DB 4; Length 477;
Best Local Similarity 41.3%; Pred. No. 2e-67;

Matches 162; Conservative 71; Mismatches 140; Indels 19; Gaps 6;

QY 2 RSLIFGATGSGISTGPDVLRKGGPEAFRTVALTGRNIRRLAEMARALKAEIAYTAHE 61
Db 81 KRISIVGSGTIGTQTDIVAE--NPDKFRVVALAAGSNVTLADQVRRFPALVAYRNE 138

QY 62 DCLPALREALAGTG--TEVAGGAQALAEADRP--ADWTMSAIVGAAGLVPGMRALKHGR 118
Db 139 SLINELKEALADLDYKLEIIPGEGVIEVARHPRPAVTVTGICAGAKPTVAALIEAGKD 198

QY 119 LALANKESLVTAAGLMTAENGATILPVDESHSAVQALAGEDTCAVEYVITTAASGP 178
Db 199 IALANKETLNGPVPVPLANKHNKXILIPADSEHSALFOCIQGLEPAGALKRIITLASGA 258

QY 179 FRDMSLERIRACTVAEAOAHPMWSMGORISIDSASMEFKALELJETREFFGFEEDRIEAV 238
Db 259 FRDMPVERLKEKVKYADALKHNNMNGKKITVDSATLFNKGLEIVEAHLFGAEYDDIEIV 318

QY 239 VHPOSIYHAMVGFCDGGLMAHLGPADMRHAIGFALNMPGR--GEVYARIDLAQIASLT 295
Db 319 IHPOSIHSMIETODSSVLAQLGMPDMRLPILYTMSPDRVPCSEVYWPRLDCKLSLT 378

QY 296 FOKPDEERFPLRLAROVMAARGISGAFAFNKAEIALDHFIAGRIAGFLDMAVAVEEYL-- 353
Db 379 FKKPDNKKYPSMDLAIYAAGRGAGTGTGVLASANKAKAEMFIDEKISYIDIFKVELTCDK 438

QY 354 --AGVSTDPLEFGKVPDALEEVLAAMDHLARRAA 383
Db 439 HRNELVTSP-----SLEEIVHYDLMAREYA 463

RESULT 4
US-09-449-335-2
Sequence 2, Application US/09449335
Patent No. 6303365
GENERAL INFORMATION:
APPLICANT: Busch, Marco
APPLICANT: Hain, Rudiger
APPLICANT: Martin, William
APPLICANT: Tietjen, Klaus
APPLICANT: Kloti, Andreas
TITLE OF INVENTION: Method of determining the activity of

TITLE OF INVENTION: 1-deoxy-D-xylose-5-phosphate reductoisomerase and

TITLE OF INVENTION: 1-deoxy-D-xylose-5-phosphate synthase

FILE REFERENCE: 2020US

CURRENT APPLICATION NUMBER: US/09/449,335

CURRENT FILING DATE: 1999-11-24

EARLIER APPLICATION NUMBER: DE 199 35 967.9

EARLIER FILING DATE: 1999-07-30

NUMBER OF SEQ ID NOS: 8

SOFTWARE: Patentln Ver. 2.1

SEQ ID NO 2

LENGTH: 477

TYPE: PRT

ORGANISM: Arabidopsis thaliana

US-09-449-335-2

[illegible]

```

Db      319  IHPOSIHSMIETODSSVLAQLGMPDMRLPLLYTMSWPDRCSEVTPRDLCKLSLT 378
Oy      296  FOKDEERFPALRLARVMAARGLSGAAPNAKKEIALDHFAGRTGLDMAVAVEETL-- 353
Db      379  FKXDNKRYTSMDLAYAGRAAGTGTGVLNANKEAVEMFIDEKISTYDLFKVVELTCDK 438
Oy      354  --AGVSTDPLEFGKVPDALEVLAMDLARRAAE---AAGLR 390
Db      439  HRNELVTSP-----SLEELVHYDLMAREYADVOJSSGAR 473

```

```

RESULT 7
US-09-491-362-2
; Sequence 2, Application US/09491362
; Patent No. 6281017
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lange, Bernd M
; TITLE OF INVENTION: 1-DEOXY-D-XYULOSE-5-PHOSPHATE REDUCTOISOMERASE, AND
; FILE REFERENCE: USUR14977
; CURRENT APPLICATION NUMBER: US/09/491,362
; EARLIER APPLICATION NUMBER: 60/118,349
; EARLIER FILING DATE: 1999-02-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Mentha piperita
US-09-491-362-2

```

```

Query Match      33.0%; Score 654.5; DB 4; Length 475;
Best Local Similarity 39.2%; Pred. No. 1.6e-58;
Matches 156; Conservative 73; Mismatches 152; Indels 17; Gaps 9;

```

```

Oy      2  RSLSTFGATGSGESTDLYMRKGGPEAFRTVALTGGNIRRLAEMARLKAELAVTAHE 61
Db      80  KPISVISTGSGIGTQTLDIYAE--NPKFRIVALAAGSNVTLADQ-KAFKPKLVSVKDE 136
Oy      62  DCLPALREALAG--TGTEVAGCAQAIAEADRP-ADWTMSAIVGAGLVPGMRALKHGR 118
Db      137  SLSELKEALAGFEDMPELIPGEGMIEVARHPDAVTVVIGVICAGLKPVAALIEGKD 196
Oy      119  LALANKESLYTAGQLMRTAOENGATILPVDESHAVFOALAGEDTACVEVITITASGP 178
Db      197  IALANKETLLAGSPFVPLAKKHNVKILPLADSEHSAIFOCIGLPEGLRRIITLASGA 256
Oy      179  FRDMSLERIRACTVAEAOAHNMGMGORISIDSASM-FNKALELITREFGFEPPRIEA 237
Db      257  FRDLPVEKLEKVKYADALKHSNMGMKNTVRLQLFENKLEVIKAYLEGAEDDIEI 316
Oy      238  VVHPOSIVHNAVVGCDGGLMAHLGPAIDMRHAIQFALNMPGR---GEVVARIDLAQIASL 294
Db      317  VHSPESTIHSNVEFQDSSVLAQLGMPDMRLPLLYTMSPEVYVSEITWPRDLCKV-DL 375
Oy      295  TFOKDERFPALRLARVMAARGLSGAAPNAKKEIALDHFAGRTGLDMAVAVEETIA 354
Db      376  FKXDNKRYTSMDLAYAGRAAGTGTGVLNANKEAVEMFIDEKISTYDLFKVVELTCD 435
Oy      355  GVSTDPLEFGKVPDALEVLAMDLARRAAE---EAAGL 389
Db      436  KHRSEM---AVSPLEELVHYDQWARYDAATVLKSAGL 470

```

```

RESULT 8
US-09-674-562-2
; Sequence 2, Application US/09874562
; Patent No. 6420159
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B

```

```

; APPLICANT: Lange, Bernd M
; TITLE OF INVENTION: 1-DEOXY-D-XYULOSE-5-PHOSPHATE REDUCTOISOMERASE, AND
; FILE REFERENCE: USUR17549
; CURRENT APPLICATION NUMBER: US/09/874,562
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 09/491,362
; PRIOR FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: 60/118,349
; PRIOR FILING DATE: 1999-02-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Mentha piperita
US-09-674-562-2

```

```

Query Match      33.0%; Score 654.5; DB 4; Length 475;
Best Local Similarity 39.2%; Pred. No. 1.6e-58;
Matches 156; Conservative 73; Mismatches 152; Indels 17; Gaps 9;

```

```

Oy      2  RSLSTFGATGSGESTDLYMRKGGPEAFRTVALTGGNIRRLAEMARLKAELAVTAHE 61
Db      80  KPISVISTGSGIGTQTLDIYAE--NPKFRIVALAAGSNVTLADQ-KAFKPKLVSVKDE 136
Oy      62  DCLPALREALAG--TGTEVAGCAQAIAEADRP-ADWTMSAIVGAGLVPGMRALKHGR 118
Db      137  SLSELKEALAGFEDMPELIPGEGMIEVARHPDAVTVVIGVICAGLKPVAALIEGKD 196
Oy      119  LALANKESLYTAGQLMRTAOENGATILPVDESHAVFOALAGEDTACVEVITITASGP 178
Db      197  IALANKETLLAGSPFVPLAKKHNVKILPLADSEHSAIFOCIGLPEGLRRIITLASGA 256
Oy      179  FRDMSLERIRACTVAEAOAHNMGMGORISIDSASM-FNKALELITREFGFEPPRIEA 237
Db      257  FRDLPVEKLEKVKYADALKHSNMGMKNTVRLQLFENKLEVIKAYLEGAEDDIEI 316
Oy      238  VVHPOSIVHNAVVGCDGGLMAHLGPAIDMRHAIQFALNMPGR---GEVVARIDLAQIASL 294
Db      317  VHSPESTIHSNVEFQDSSVLAQLGMPDMRLPLLYTMSPEVYVSEITWPRDLCKV-DL 375
Oy      295  TFOKDERFPALRLARVMAARGLSGAAPNAKKEIALDHFAGRTGLDMAVAVEETIA 354
Db      376  FKXDNKRYTSMDLAYAGRAAGTGTGVLNANKEAVEMFIDEKISTYDLFKVVELTCD 435
Oy      355  GVSTDPLEFGKVPDALEVLAMDLARRAAE---EAAGL 389
Db      436  KHRSEM---AVSPLEELVHYDQWARYDAATVLKSAGL 470

```

```

RESULT 9
US-09-368-590-2
; Sequence 2, Application US/09368590
; Patent No. 6187563
; GENERAL INFORMATION:
; APPLICANT: Solimena, Michele
; TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
; FILE REFERENCE: 101918-200 (OCR-941)
; CURRENT APPLICATION NUMBER: US/09/368,590
; CURRENT FILING DATE: 1999-08-04
; EARLIER APPLICATION NUMBER: 60/095,657
; EARLIER FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2293
; TYPE: PRT
; ORGANISM: Human
US-09-368-590-2

```

```

Query Match      7.0%; Score 138.5; DB 4; Length 2293;

```


Best Local Similarity 25.6%; Pred. No. 0.00012;
Matches 103; Conservative 43; Mismatches 138; Indels 119; Gaps 23;

```

QY 42 RRLAARALKALTAEDLCPLAREALAGTGVAGAAIA-----86
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 507 RRLARHRLATGE--VEARGPVSGRLRLATLGGASGAPLVLAQVREAEOLFAEV 564
QY 87 -EADRPADMTSAIVGAAGLVPMGRALKHRTIALANKESLVTAGOLMRTAENGATI 145
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 565 TEVAALRRWLMDALA-----YRAMEGVHACELWIGKEQWL-----LSMR-----606
QY 146 LPVDSEHSAY----FQALGEDTACVERVI-----ITASGPRDMSLERIRACTVA 193
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 607 VPDSDLDVEVVOHREFESLQEMNSLMGRVLDVHTVQELVEGHP-----SSDEVRSCT---659
QY 194 EAQAHNMMSGORISIDSAMFNKALELTETREFEFEDRIEAVY----HPOSI---VHA 247
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 660 --QDHLN-----SRMRTVELVEORR-----EEMSAVLLENHVELEVAEVR 699
QY 248 MVGFCDGLMAHLGPMRHAIA-----GFALNMPGRG--EVPARIDLAQIASLTFQKPD 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 700 QY-----REKRRAVESAPRAGALQWRLSGLEALQALEPQOALEEAL 746
QY 301 EERFPA-----LRLADYMAAR--GLSGAAPNAKKEIA-----LDHTAGRIEIDMAAYE 350
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 747 AERFPAQAAAXLHQAEELCAEWGALASAAQCEAAVAAAGRLQRFHLDLDAFLDWLVRAQ 806
QY 351 ETLGAVSTDLFGKVPDALAEVYLAHMDHLARRAA--EEAAGLRQ 392
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 807 EA-AGSGEPL-----PNSLEADAL--LARHAAKKEVDOR 842

```

RESULT 10
US-09-105-537-2

```

; Sequence 2, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and plikromycin
; FILE REFERENCE: 600,438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 5215
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-105-537-2

```

Query Match 5.8%; Score 115; DB 4; Length 5215;
Best Local Similarity 20.2%; Pred. No. 0.11;
Matches 108; Conservative 52; Mismatches 162; Indels 212; Gaps 21;

```

QY 28 EAFRTVALTGRNIRRLAMARALKAEIA-----VTAHEDCLPLAREALAGTGEVAG 80
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1509 QAVRLTEFLDGRDVTVALDGLSLARTFENKALTTTATRELLAGDAL--GRGQATG 1567
QY 81 -----GAQOIA-----EADRPADMTSAIV 101
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1568 IYTGEPARGRTAFLFTGGGAQRVAMGEELRAHPVFAALDVTVAALDRHIDRLREIV 1627
QY 102 GAA-----GLVP-----GMRALKH--GRTLALANKE 125
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1628 AAGEELDLATYQPALFAFEVALFRLEHNGLVPRDLTGHSVGEIAAAHVAGVSLDDAA 1687
QY 126 SLVTGOLMRTAENGATILPVDSEHSAYFQALGEDTACVERVIITASGFF-----179
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1688 RLVTARGRLMOSAREGGA--MIAVQAGEAEVYESLKGYE-----GRVAAVAVNGPTAVVSG 1742

```

```

QY 180 -RDMSLERIRACTVAEQAHPMWS--MGORI-----SIDSAMFNKALELTETREFEFG 230
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1743 DAD-AAEELRAV-----WAGRGRTRLRVSHAHSPHMDVDLDEPLRVAGLELF 1791
QY 231 EPRDIAVYVPOSIYHAMV-----GECGGGLMAHLCP 262
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1792 EEPRIPIV--STVTGALTSGELTSPAYVWDQIRRPVFEFLDAVRTLAQADATVLEIOP 1848
QY 263 ADMRHAIGFALNMPGRGEPVVARIDLAQIASLTFQKPDDE-----302
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1849 DAVLTALAEALAPGT--DAPDAR--DVTVPPLRAGRPEPBTTLAAGLATYHVHAPLRAS 1906
QY 303 -----RFPALRLADYMAARG--LSGA 322
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1907 FFPDGRRTDLPYAFRRHXYWLTPEARDTARALGPDRAHPLITTTVEVAGDGYLLTR 1966
QY 323 AFNAKEIALDHTIAGRI-----GFLDMAAVYEETLAGYSTDLFGKVDAL 371
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1967 LSLTDQPMWLADHVNGAVLLPATFLELALAGDHGAVREELTLEAPLVLE 2020

```

RESULT 11
US-09-335-409-5

```

; Sequence 5, Application US/09335409
; Patent No. 6121029
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goelach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/335,409
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7257
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-335-409-5

```

Query Match 5.7%; Score 113.5; DB 3; Length 7257;
Best Local Similarity 24.0%; Pred. No. 0.26;
Matches 95; Conservative 42; Mismatches 136; Indels 123; Gaps 19;

```

QY 4 LSIIFGATGSGESTFDLVNRKGGPEAFRTVALTGRNIRRLAMARALKAEIA-----VT 58
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5510 VSGFGVSG-----TNAHVVEEAPFAAPAAARSALFVLSAKSAALDAQARLSAHV 5565
QY 59 AHEDC--LPALREALAGTGEV-----AGCAOIAEADRPADMTSAIVGAAGLVPM 110
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5566 AHEPELIGDLASLATTRSPMYRLAVAATSRALSALD-----TAAQGGAPPA 5615
QY 111 RLKHGRTLANKESLVTAGOLMRTAENGATILPVDSEHSAYFQALGED-----163
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5616 AARGHAST--GSAPKVVFVPPG-----GSOWLGMOCKLLSEPPVFRDAL 5658
QY 164 TACVERVIITASGPRDMSLERIRACTVAEQAHPMWSGORISIDSAMFNKALELT 223
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5659 SAC--DRAIQANG-----WSL-----LAEALADDTTSGLDIVVQALPAIEVALSA 5705
QY 224 TREFFGEFPRDIEAVYVPOSIYHAMVGFCDGLMAHLGPMRHAIGFALNMPGCEVPV 283
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5706 LWRSMGVPEPDVAV-----VGHSM-----GEVAA 5727
QY 284 ARIDLAQIASLTFQKPDDEERFPALRLADYMAARGLSGAAPNAKKEIALDHTIAGRI 343
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5728 AHV--AGALSL-----EDAVAIT--ICRSLLLRKISGGGEAAVVELSLAEABALLGYE 5777
QY 344 DMAAVVEETLAGYSTDP-----LFGKVPDALAEVYLA 375

```

Db 5778 DRLSV-----AVSNSPRSTVLAGE-PALALAEVLAI 5806

RESULT 12

US-09-568-102-5

Sequence 5, Application US/09568102

Patent No. 6346404

GENERAL INFORMATION:

APPLICANT: Schupp, Thomas

APPLICANT: Ligon, James

APPLICANT: Molnar, Istvan

APPLICANT: Zirkle, Ross

APPLICANT: Goerlach, Joern

TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES

FILE REFERENCE: 4-30582A

CURRENT APPLICATION NUMBER: US/09/568,102

CURRENT FILING DATE: 2000-05-10

PRIOR APPLICATION NUMBER: 09/335,409

PRIOR FILING DATE: 1999-06-17

NUMBER OF SEQ ID NOS: 30

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 5

LENGTH: 7257

TYPE: PRT

ORGANISM: Sorangium cellulosum

US-09-568-102-5

Query Match

Best Local Similarity 5.7%; Score 113.5; DB 4; Length 7257;

Matches 95; Conservative 42; Mismatches 136; Indels 123; Gaps 19;

QY 4 LSIFGATGSGESTFDLVRKRGPEAFRTVALTGGRNIRRLAEMARALKAEIA-----VT 58

Db 5510 VSSFGVSG-----TNAHVYLEAPAAAFAPAPARSASELFVLSAKSAALDQAARLSAHV 5565

QY 59 AHEDC-LPALREALAGTGEV-----AGGAQAIAEADRPADMTWSAIVGAAGLVPGM 110

Db 5566 AHPGLGDLAFSLATRSPTVRLAANAATSRALSAALD-----TAAQQAAPPA 5615

QY 111 RALKHGRTLANKESLVTAGQLMRTAENGATILPVDSHSAVFOALAGED----- 163

Db 5616 AARGHAST-GSAPRVVFEPFQ-----GSGWLMGOKLSEEPVFRDAL 5658

QY 164 TACVERVITTAGCGPFDMSLERIRACTVAEAQHPNMSGQRISIDSASFNKALELIE 223

Db 5659 SAC-DRAIQAEAG-----WSL-----LAELADETTSQGRIDVQPAFAIEVALSA 5705

QY 224 TREFFGEPPRIEAVHPQSIVHAMVGFCDGGLMAHLGPRADMRHAIGFALNMPGRGEVPV 283

Db 5706 LMRSMGVEPPDAV-----VGHSM-----GEVAA 5727

QY 284 ARIDLAQIASLTFQKPDREPPALRLARDVMAARGLSGAAPNAKETAIDHFIAGRIGFL 343

Db 5728 AHV--AGALSL-----EDAVAI-ICRSSLRLRISGCGMAVVELSLAEAEALLGIE 5777

QY 344 DMAAVEETLAGVSTDP-----LFGKVPDALEEVLA 375

Db 5778 DRLSV-----AVSNSPRSTVLAGE-PALALAEVLAI 5806

RESULT 13

US-09-567-969-5

Sequence 5, Application US/09567969

Patent No. 6355457

GENERAL INFORMATION:

APPLICANT: Schupp, Thomas

APPLICANT: Ligon, James

APPLICANT: Molnar, Istvan

APPLICANT: Zirkle, Ross

APPLICANT: Cyr, Devon

APPLICANT: Goerlach, Joern

TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES

FILE REFERENCE: 4-30582A

CURRENT APPLICATION NUMBER: US/09/567,969

CURRENT FILING DATE: 2000-05-10

PRIOR APPLICATION NUMBER: 09/335,409

PRIOR FILING DATE: 1999-06-17

NUMBER OF SEQ ID NOS: 30

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 5

LENGTH: 7257

TYPE: PRT

ORGANISM: Sorangium cellulosum

US-09-567-969-5

Query Match

Best Local Similarity 5.7%; Score 113.5; DB 4; Length 7257;

Matches 95; Conservative 42; Mismatches 136; Indels 123; Gaps 19;

QY 4 LSIFGATGSGESTFDLVRKRGPEAFRTVALTGGRNIRRLAEMARALKAEIA-----VT 58

Db 5510 VSSFGVSG-----TNAHVYLEAPAAAFAPAPARSASELFVLSAKSAALDQAARLSAHV 5565

QY 59 AHEDC-LPALREALAGTGEV-----AGGAQAIAEADRPADMTWSAIVGAAGLVPGM 110

Db 5566 AHPGLGDLAFSLATRSPTVRLAANAATSRALSAALD-----TAAQQAAPPA 5615

QY 111 RALKHGRTLANKESLVTAGQLMRTAENGATILPVDSHSAVFOALAGED----- 163

Db 5616 AARGHAST-GSAPRVVFEPFQ-----GSGWLMGOKLSEEPVFRDAL 5658

QY 164 TACVERVITTAGCGPFDMSLERIRACTVAEAQHPNMSGQRISIDSASFNKALELIE 223

Db 5659 SAC-DRAIQAEAG-----WSL-----LAELADETTSQGRIDVQPAFAIEVALSA 5705

QY 224 TREFFGEPPRIEAVHPQSIVHAMVGFCDGGLMAHLGPRADMRHAIGFALNMPGRGEVPV 283

Db 5706 LMRSMGVEPPDAV-----VGHSM-----GEVAA 5727

QY 284 ARIDLAQIASLTFQKPDREPPALRLARDVMAARGLSGAAPNAKETAIDHFIAGRIGFL 343

Db 5728 AHV--AGALSL-----EDAVAI-ICRSSLRLRISGCGMAVVELSLAEAEALLGIE 5777

QY 344 DMAAVEETLAGVSTDP-----LFGKVPDALEEVLA 375

Db 5778 DRLSV-----AVSNSPRSTVLAGE-PALALAEVLAI 5806

RESULT 14

US-09-568-480-5

Sequence 5, Application US/09568480

Patent No. 6355458

GENERAL INFORMATION:

APPLICANT: Schupp, Thomas

APPLICANT: Ligon, James

APPLICANT: Molnar, Istvan

APPLICANT: Zirkle, Ross

APPLICANT: Cyr, Devon

APPLICANT: Goerlach, Joern

TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES

FILE REFERENCE: 4-30582A

CURRENT APPLICATION NUMBER: US/09/568,480

CURRENT FILING DATE: 2000-05-10

PRIOR APPLICATION NUMBER: 09/335,409

PRIOR FILING DATE: 1999-06-17

NUMBER OF SEQ ID NOS: 30

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 5

LENGTH: 7257

TYPE: PRT

ORGANISM: Sorangium cellulosum

US-09-568-480-5

Query Match

Best Local Similarity 5.7%; Score 113.5; DB 4; Length 7257;

Best Local Similarity 24.0%; Pred. No. 0.26;
Matches 95; Conservative 42; Mismatches 136; Indels 123; Gaps 19;

```
OY 4 LSIFGATSGESTFDLMRKGGPEAFRTVALTGRNIRRLAEMARKAEIA-----VT 58
Db 5510 VSSFGVSG-----TNAHVLEEPAPAAAFAPAAARSALFVLSAKSAALDAQAARLSAHV 5565
OY 59 AHEDC-LPALREALAGTGEV-----AGGAQAIAEADRPADMTSAIYGAAGLVPGM 110
Db 5566 AHPELGIDDLAFSLATTSPTMYRLAVNATSRALSAID-----TAAGQAPPA 5615
OY 111 RALKHGRLLALANKESLVTAGQLMRTAENGATLLPYDSESAVFOALAGED----- 163
Db 5616 AARGHAST-GSAPKVVFFVPGQ-----GSQWLGKQKLLSEEPFRDAL 5658
OY 164 TACVERVITITASGPFPRDMSLERIACVTAEOAHPPNMSGORISIDSASFENKALELIE 223
Db 5659 SAC-DRAIQAEAG-----WST-----LAEIADETTSQLGRIDVYQPALFAIEVALSA 5705
OY 224 TREFFGEPPDRIEAVVHPQSIYHAWGFCGGLMAHLGPADMRHAIGFALMMPGRGEVPV 283
Db 5706 LMRSGVPEPDV-----VGHSM-----GEVAA 5727
OY 284 ARIDIAQIASLTFOKPDERFPALRLARDVMAARGLSGAAPNAKEIALDHFIAGRIGFL 343
Db 5728 AHV--AGALSL-----EDAVAI-ICRRSLRLRISGGEWAVVELSLAEAEALLGYE 5777
OY 344 DMAAVEETLAGVSTDP-----LFGKVPDALEEVILAM 375
Db 5778 DRLSY-----AVSNSPRSTVLAGE-PAALAEVLAI 5806
```

RESULT 15

```
US-09-568-486-5
; Sequence 5, Application US/09568486
; Patent No. 6355459
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligdon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zitzke, Ross
; APPLICANT: Cyt, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,486
; PRIOR FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 5
; LENGTH: 7257
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-568-486-5
```

Query Match 5.7%; Score 113.5; DB 4; Length 7257;

Best Local Similarity 24.0%; Pred. No. 0.26;
Matches 95; Conservative 42; Mismatches 136; Indels 123; Gaps 19;

```
OY 4 LSIFGATSGESTFDLMRKGGPEAFRTVALTGRNIRRLAEMARKAEIA-----VT 58
Db 5510 VSSFGVSG-----TNAHVLEEPAPAAAFAPAAARSALFVLSAKSAALDAQAARLSAHV 5565
OY 59 AHEDC-LPALREALAGTGEV-----AGGAQAIAEADRPADMTSAIYGAAGLVPGM 110
Db 5566 AHPELGIDDLAFSLATTSPTMYRLAVNATSRALSAID-----TAAGQAPPA 5615
OY 111 RALKHGRLLALANKESLVTAGQLMRTAENGATLLPYDSESAVFOALAGED----- 163
Db 5616 AARGHAST-GSAPKVVFFVPGQ-----GSQWLGKQKLLSEEPFRDAL 5658
```

```
OY 164 TACVERVITITASGPFPRDMSLERIACVTAEOAHPPNMSGORISIDSASFENKALELIE 223
Db 5659 SAC-DRAIQAEAG-----WST-----LAEIADETTSQLGRIDVYQPALFAIEVALSA 5705
OY 224 TREFFGEPPDRIEAVVHPQSIYHAWGFCGGLMAHLGPADMRHAIGFALMMPGRGEVPV 283
Db 5706 LMRSGVPEPDV-----VGHSM-----GEVAA 5727
OY 284 ARIDIAQIASLTFOKPDERFPALRLARDVMAARGLSGAAPNAKEIALDHFIAGRIGFL 343
Db 5728 AHV--AGALSL-----EDAVAI-ICRRSLRLRISGGEWAVVELSLAEAEALLGYE 5777
OY 344 DMAAVEETLAGVSTDP-----LFGKVPDALEEVILAM 375
Db 5778 DRLSY-----AVSNSPRSTVLAGE-PAALAEVLAI 5806
```

Search completed: April 15, 2003, 14:00:29
Job time : 37 secs

	RESULT
AAV52839	ID
AAV52839 standard; Protein; 394 AA.	
AAV52839;	
26-JAN-2000 (first entry)	
Rhodobacter sphaeroides protein sequence SEQ ID NO:30.	
Isoprenoid; microorganism; detection; antibacterial; herbicide; heart disease; osteoporosis; haemostasis; cancer; immunopotentialion health food; antitumor coating; farnesyl pyrolinic acid; pyruvic 1-deoxy-D-xylulose-5-phosphate; glyceraldehyde-3-phosphate; 2-C-methyl-D-erythreitol-4-phosphate.	
Rhodobacter sphaeroides.	
WO9953071-A1.	
21-OCT-1999.	
14-APR-1999; 99WO-JP01987.	
14-APR-1998; 98JP-0103101. 05-AUG-1998; 98JP-0221910. 15-FEB-1999; 99JP-0035739.	
(KYOWA HAKKO KOGYO KK.	
Miyake K, Hashimoto S, Motoyama H, Ozaki A, Seto H, Kuzuyama T, Takamashi S;	

DR WP1: 1999-620434/53.
DR N-PSDB: AAZ33164.
PT Preparation of recombinant isoprenoid compounds useful for treatment of
PT heart diseases, osteoporosis and hemostasis, preventing cancer and
PT immunopotentialiation
PS Claim 6; Page 130-132; 145pp; Japanese.
XX
XX The present invention describes the preparation of an isoprenoid
CC compound comprising using at least 1 DNA e.g. encoding proteins which
CC elevate the efficiency of the synthesis or DNA encoding a farnesyl
CC pyrolytic producing enzyme. The method of preparation of an isoprenoid
CC compound comprises using at least 1 DNA, a vector, cloned cells, their
CC derived recombinant DNAs or transformed products in a culture system and
CC extracting the isoprenoid accumulated in the medium. The DNA encodes at
CC least 1 of the following: (a) a compound for activating or catalyzing
CC the production of 1-deoxy-D-xylulose-5-phosphate from pyruvic acid and
CC glyceraldehyde-3-phosphate; (b) an enzyme producing farnesyl pyrolytic
CC acid; (c) a protein which elevates the efficiency of synthesis of
CC isoprenoid compounds and comprises a 3 or 4 amino acid sequence
CC optionally with 1 or more of the amino acids being deleted or
CC substituted or an additional amino acid being inserted; (d) a protein
CC which activates or catalyzes the production of 2-C-methyl-D-erythritol-
CC 4-phosphate from 1-deoxy-D-xylulose-5-phosphate; or (e) a protein which
CC activates a target compound or reaction and is a string end or hybrid of
CC the DNA encoded in (a)-(e). Isoprenoid compounds are useful in drugs
CC (e.g. for the treatment of heart diseases, osteoporosis and hemostasis,
CC for preventing cancer and as immunopotentializers), health foods and
CC antifouling coatings. The isoprenoids also inhibit enzymatic reactions
CC on the non-mevalonate pathway and can be used as antibacterials and
CC herbicides. The present sequence is used in the exemplification of
CC the present invention.
XX
SQ Sequence 394 AA;
Query Match 100.0%; Score 1984; DB 20; Length 394;
Best Local Similarity 100.0%; Pred. No. 1.8e-193;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRSLSIFGATGSGESTFDLMRRKGPEAFRTVALTGGRNIRRLAEMARALKAELAVTAH 60
Db 1 MRSLSIFGATGSGESTFDLMRRKGPEAFRTVALTGGRNIRRLAEMARALKAELAVTAH 60
QY 61 EDCPLALREALAGTGEVAGGAQAIAEADRPADMTSAIYGAGLVGMRALKHGRTLA 120
Db 61 EDCPLALREALAGTGEVAGGAQAIAEADRPADMTSAIYGAGLVGMRALKHGRTLA 120
QY 121 LANKESLVTAGQLMRTAENGATILLPVDSHSAVFQALAGEDTRACVERVITITASGGPFR 180
Db 121 LANKESLVTAGQLMRTAENGATILLPVDSHSAVFQALAGEDTRACVERVITITASGGPFR 180
QY 181 DWSLERIRACTVAEAQAHPNMSMGORISIDSASFNKALELIEFREFFGFPDRIEAVYH 240
Db 181 DWSLERIRACTVAEAQAHPNMSMGORISIDSASFNKALELIEFREFFGFPDRIEAVYH 240
QY 241 POSTIVHAMVGFCDGGLMAHIGPADMRHAIGFALWMPGGEVVARIDLAQIASLTFOKPD 300
Db 241 POSTIVHAMVGFCDGGLMAHIGPADMRHAIGFALWMPGGEVVARIDLAQIASLTFOKPD 300
QY 301 EERRPALRLADVMAAAGLSGAAPNAKKEIALDHFIAGRIGFLDMAAVVEETLAGVSTDP 360
Db 301 EERRPALRLADVMAAAGLSGAAPNAKKEIALDHFIAGRIGFLDMAAVVEETLAGVSTDP 360
QY 361 LFGKVPDALEEVILAMDHILARRAAEEAAGLRQOKR 394
Db 361 LFGKVPDALEEVILAMDHILARRAAEEAAGLRQOKR 394
RESULT 2
AAO21875
ID AAO21875 standard; Protein; 386 AA.
XX

AC AAO21875;
XX
XX 13-SEP-2002 (first entry)
DE Isoprenoid related protein sequence SEQ ID No 97.
XX
XX Isoprenoid; CoQ(10); 1-deoxyxylulose-5-phosphate synthase; DXS; DDS;
XX decaprenyl diphosphate synthase.
XX Spilngomonas treperei.
XX WO200226933-A2.
XX
XX 04-APR-2002.
XX
XX 28-SEP-2001; 2001WO-US30328.
XX
XX 29-SEP-2000; 2000US-236580P.
XX
XX (CRGI) CARGILL INC.
XX Gokarn R, Jessen H, Zidwick MJ;
XX
XX WP1: 2002-416480/44.
XX
XX Substantially pure polypeptides having e.g.,
PT 1-deoxyxylulose-5-phosphate synthase activity, useful for the
PT production of isoprenoids, especially CoQ(10)
XX
XX Disclosure; Fig 30; 246pp; English.
XX
XX The invention relates to methods and materials for the production of
CC isoprenoids. More particularly the invention provides isolated nucleic
CC acids, substantially pure polypeptides, host cells, and methods for
CC producing various isoprenoid compounds. The polypeptides are useful for
CC the production of isoprenoids, especially CoQ(10). Expressing the pure
CC polypeptides, which has 1-deoxyxylulose-5-phosphate synthase (DXS)
CC activity or decaprenyl diphosphate synthase (DDS) activity, is useful for
CC increasing production of CoQ(10) in a cell having endogenous DDS
CC activity. This sequence represents a protein relating to the isoprenoid
CC production of the invention.
XX
SQ Sequence 386 AA;
Query Match 46.3%; Score 918.5; DB 23; Length 386;
Best Local Similarity 50.4%; Pred. No. 8e-85;
Matches 194; Conservative 62; Mismatches 122; Indels 7; Gaps 3;
QY 1 MRSLSIFGATGSGESTFDLMRRKGPEAFRTVALTGGRNIRRLAEMARALKAELAVTAH 60
Db 2 VKRVTVLGATGSGVSTLTLIER--NPAFEVVALTANCDDVEKLAAAIIRTRACAVVAD 59
QY 61 EDCPLALREALAGTGEVAGGAQAIAEADRPADMTSAIYGAGLVGMRALKHGRTLA 120
Db 61 EDCPLALREALAGTGEVAGGAQAIAEADRPADMTSAIYGAGLVGMRALKHGRTLA 120
QY 121 LANKESLVTAGQLMRTAENGATILLPVDSHSAVFQALAGEDTRACVERVITITASGGPFR 180
Db 121 LANKESLVTAGQLMRTAENGATILLPVDSHSAVFQALAGEDTRACVERVITITASGGPFR 180
QY 181 DWSLERIRACTVAEAQAHPNMSMGORISIDSASFNKALELIEFREFFGFPDRIEAVYH 240
Db 181 DWSLERIRACTVAEAQAHPNMSMGORISIDSASFNKALELIEFREFFGFPDRIEAVYH 240
QY 241 POSTIVHAMVGFCDGGLMAHIGPADMRHAIGFALWMPGGEVVARIDLAQIASLTFOKPD 300
Db 241 POSTIVHAMVGFCDGGLMAHIGPADMRHAIGFALWMPGGEVVARIDLAQIASLTFOKPD 300
QY 301 EERRPALRLADVMAAAGLSGAAPNAKKEIALDHFIAGRIGFLDMAAVVEETLAGVSTDP 360
Db 301 EERRPALRLADVMAAAGLSGAAPNAKKEIALDHFIAGRIGFLDMAAVVEETLAGVSTDP 360
QY 361 LFGKVPDALEEVILAMDHILARRAAEE 385
Db 361 LFGKVPDALEEVILAMDHILARRAAEE 385

Db 358 ---AAPTLDVLAIDAFAALYAAE 379

RESULT 3

AAO21879 standard; Protein; 388 AA.

AAO21879;

13-SEP-2002 (first entry)

Isoprenoid related protein sequence SEQ ID No 101.

Isoprenoid; CoQ(10); 1-deoxyxylulose-5-phosphate synthase; DXS; DDS; decaprenyl diphosphate synthase.

Zymonas mobilis.

WO200226933-A2.

04-APR-2002.

28-SEP-2001; 2001WO-US30328.

29-SEP-2000; 2000US-236580P.

(CRGI) CARGILL INC.

Gokarn R, Jessen H, Zidwick MJ;

WPI; 2002-416480/44.

Substantially pure polypeptides having e.g., 1-deoxyxylulose-5-phosphate synthase activity, useful for the production of isoprenoids, especially CoQ(10)

Disclosure; Fig 27; 246pp; English.

The invention relates to methods and materials for the production of isoprenoids. More particularly the invention provides isolated nucleic acids, substantially pure polypeptides, host cells, and methods for producing various isoprenoid compounds. The polypeptides are useful for the production of isoprenoids, especially CoQ(10). Expressing the pure polypeptides, which has 1-deoxyxylulose-5-phosphate synthase (DXS) activity or decaprenyl diphosphate synthase (DDS) activity, is useful for increasing production of CoQ(10) in a cell having endogenous DDS activity. This sequence represents a protein relating to the isoprenoid production of the invention.

Sequence 388 AA;

Query Match 44.2%; Score 876; DB 23; Length 388;

Best Local Similarity 46.9%; Pred. No. 1.7e-80; Matches 183; Conservative 67; Mismatches 130; Indels 10; Gaps 3;

2 RLSITGATGSGESTFEDLMRKSGEAPRTVALTNGRNIRLAEMARLAKELAVTAHE 61

5 RVTYVAGATGSGHSTLIER--NLDROYALVTANRWKDLADAKTKNRAYIADP 62

62 DCLPALREALAGTGVAGCAQAIAEADRPADMTSAIVGAAGLVPGMRALKHGTTLA 121

63 SLYNDLKEALAGSSVEAAGADALVEAAMGADMTMAIICGGLKATLAIIRKGTVAL 122

122 ANKESLVTAAGOLMRTAENGATILPVDSEHSVAFQALAGEPTACVERVITITASGGPFD 181

123 ANKESLVTAAGOLMRTAENGATILPVDSEHSVAFQALAGEPTACVERVITITASGGPFD 182

182 WSLERIRACTVAAEQAHPMWMSGORISIDSMFNKALELIERFEFFGEPPDRIEAVHP 241

183 TSLAEAVATYTPERAVOHPMWSMAKISIDSATMMNKGLIELIEXHLPFPLEKFELTVHP 242

242 OSTIVHAMVGCDSGLMAHLGIPADMRAHIGFALNWPGRGEVPAVIDLAQIASLTFOKPDE 301

Db 243 OSYIHSWVEYLDGSIQAQSGPDMRPTIGHTLAMPFRMETPAESIDFTLRQMDPEADPY 302

302 EREPALRLARDVMAANGSGAFAFNAAKETALDHFHFGIFLDMAAVVEETLAGSTDP 361

303 EREPALTLAMESIKSGAPPAVMNANETAVAFDCKIGFLDIAXIVERKL-----DHY 357

362 FGKVPDALEEVLMADHLARRAEAGLRO 391

358 TPATPSSLEDVFAIDNEAR---IQAAALME 384

RESULT 4

AAO21883 standard; Protein; 388 AA.

AAO21883;

13-SEP-2002 (first entry)

Isoprenoid related protein sequence SEQ ID No 116.

Isoprenoid; CoQ(10); 1-deoxyxylulose-5-phosphate synthase; DXS; DDS; decaprenyl diphosphate synthase.

Zymonas mobilis.

WO200226933-A2.

04-APR-2002.

28-SEP-2001; 2001WO-US30328.

29-SEP-2000; 2000US-236580P.

(CRGI) CARGILL INC.

Gokarn R, Jessen H, Zidwick MJ;

WPI; 2002-416480/44.

Substantially pure polypeptides having e.g., 1-deoxyxylulose-5-phosphate synthase activity, useful for the production of isoprenoids, especially CoQ(10)

Disclosure; Fig 32; 246pp; English.

The invention relates to methods and materials for the production of isoprenoids. More particularly the invention provides isolated nucleic acids, substantially pure polypeptides, host cells, and methods for producing various isoprenoid compounds. The polypeptides are useful for the production of isoprenoids, especially CoQ(10). Expressing the pure polypeptides, which has 1-deoxyxylulose-5-phosphate synthase (DXS) activity or decaprenyl diphosphate synthase (DDS) activity, is useful for increasing production of CoQ(10) in a cell having endogenous DDS activity. This sequence represents a protein relating to the isoprenoid production of the invention.

Sequence 388 AA;

Query Match 44.2%; Score 876; DB 23; Length 388;

Best Local Similarity 46.9%; Pred. No. 1.7e-80; Matches 183; Conservative 67; Mismatches 130; Indels 10; Gaps 3;

2 RLSITGATGSGESTFEDLMRKSGEAPRTVALTNGRNIRLAEMARLAKELAVTAHE 61

5 RVTYVAGATGSGHSTLIER--NLDROYALVTANRWKDLADAKTKNRAYIADP 62

62 DCLPALREALAGTGVAGCAQAIAEADRPADMTSAIVGAAGLVPGMRALKHGTTLA 121

63 SLYNDLKEALAGSSVEAAGADALVEAAMGADMTMAIICGGLKATLAIIRKGTVAL 122

122 ANKESLVTAAGOLMRTAENGATILPVDSEHSVAFQALAGEPTACVERVITITASGGPFD 181

Db 123 ANKESLVSAGLMDADVAHEGCTLLPVDSEHNALFQCEPHNNRDYVRRIITTAGSGPRT 182
 QY 182 WSLERIRACTVAEQAHPNMSGORISIDSASFNKALELIEBFEEDREAVVHP 241
 Db 183 TSLAEATVTERAVQHPNMSGAKISIDSATMANKLELEAVHFOIPLEKEIILVHP 242
 QY 242 QSIYHAWGFCDDGLMAHLGPADNRHAIIGFALNMPGRGEVVARIDLAQIASLTFQKDE 301
 Db 243 QSVIHSWVEYIDGSIILAQIGSPDMRTPIGHTLAMPKREMTAESLIDFTKLQMPDEADY 302
 QY 302 ERFAPLRLARVMAARGISGAFAFNAAKEIALDHTIAGRIGLDMAAVVEETLAVSDPL 361
 Db 303 ERFAPLRLAMESIKSGARPAVMNNAEIAVAEFLDKKIGLIDAKIVETL-----DHY 357
 QY 362 FGKVPDALEEVLAAMDHLARRAAEEAQLRQ 391
 Db 358 TPATPSSLEDFALDNEAR---IQAAALME 384
 RESULT 5
 ABG61582
 ID ABG61582 standard; protein; 394 AA.
 AC ABG61582;
 XX 27-AUG-2002 (first entry)
 DE High growth methanotrophic bacterial strain: polypeptide #32.
 XX High growth methanotrophic bacterial strain; C1 carbon substrate; enzyme:
 KM methane; methanol; Embden-Meyerhof carbon flux pathway; 16S RNA;
 KM pyrophosphate dependent phosphofructokinase; nitrogen-containing compound;
 KM ammonia; nitrate; nitrite; nitrogen; pigment; oxygen; landfill;
 KM methane-containing environment; waste water treatment system; isoprenoid;
 KM nitrous oxide; terpenoid; animal feed; carotenoid; exopolysaccharide.
 XX Methylobionas 16a.
 OS WO200220728-A2.
 PN 14-MAR-2002.
 PD 28-AUG-2001; 2001WO-US26827.
 PE 01-SEP-2000; 2000US-229858P.
 PR (DUPO) DU PONT DE NEMOURS & CO E I.
 PA Koffas M, Odom JM, Schenzle A;
 XX Koffas M, Odom JM, Schenzle A;
 PI WPI; 2002-452200/48.
 DR N-PSDB; ABK83261.
 XX
 PT New high growth methanotrophic bacterial strain, useful for producing
 PT single cell proteins, grows on a C1 carbon substrate, and comprises a
 PT functional gene encoding in Embden-Meyerhof carbon pathway
 XX
 PS Claim 11; Page 141-142; 157pp; English.
 CC The invention relates to a high growth methanotrophic bacterial strain,
 CC which grows on a C1 carbon substrate e.g. methane and methanol, and
 CC comprises a functional Embden-Meyerhof carbon flux pathway comprising a
 CC gene coding a pyrophosphate dependent phosphofructokinase enzyme or a 16S
 CC RNA. The bacterial strain is useful for the production of single cell
 CC protein and for the biotransformation of a nitrogen-containing compound,
 CC e.g. ammonia, nitrate or nitrogen. It is also useful for the
 CC production of a feed product comprising a protein, carbohydrates and a
 CC pigment and for reducing oxygen demand, for removing nitrates and
 CC nitrates in methane-containing environments such as landfills, waste
 CC water treatment systems or anywhere that methane, oxygen and nitrates are
 CC present. The bacterial strain of the invention can be used as a
 CC denitrifying agent for the conversion of nitrate or nitrite to nitrous

CC oxide with methane or methanol as a carbon source. It is also used in the
 CC production of biomass including proteins, carbohydrates and a wide
 CC variety of pigments (particularly for isoprenoid pigments for the
 CC purpose of generating animal feeds). In production of terpenoid and
 CC carotenoid compounds, useful as pigments and as monomers in polymeric
 CC materials and in production of exopolysaccharides at high levels.
 CC Sequences ABG61551-ABG61590 represent high growth methanotrophic
 CC bacterial strain proteins of the invention.
 XX
 SQ Sequence 394 AA;
 Query Match 39.28; Score 778; DB 23; Length 394;
 Best Local Similarity 41.68; Pred. No. 1.7e-70;
 Matches 164; Conservative 75; Mismatches 139; Indels 16; Gaps 5;
 QY 1 MRSSTIFGATSGISGSEFEDLYMRKSGPEAFRTVALTGGRRNRLAEAMARKLAETAAH 60
 Db 1 MKGICITGATSGISGVSTLDVVARH--PDKYOVVALTANGNDALYEGLAHHPYAVVVM 58
 QY 61 EDCPLALREALAGT---GTEVAGGAQAIAEAAD--RPADWTSAIVGAAGLVPGMRALKHG 116
 Db 59 ESKVAEFKQRIASPVDIKVLSGSEALQVATLENDVTWAAIVGAAGLPTLAARKAG 118
 QY 117 RTLALANKESLVTAGQLMRTAQENGATILPVDSHSAVFOAL-----AGEDTACVERVI 171
 Db 119 KTVLLANKKALVMGQIFMQAVSDSGAVLLPIDEHNNAIFQCMGPGYTPGHTAKQARRIL 178
 QY 172 ITASGPFPRDMSLERIRACTVAEQAHPNMSGORISIDSASFNKALELIEBFEEDRE 231
 Db 179 LTASGPFPRRPIETLSSVTPDOAVAHPKMDGKISVDSATMANKLELIEACLELNE 238
 QY 232 PDRIEAVVHPOSIYHAWGFCDDGLMAHLGPADNRHAIIGFALNMPGRGEVVARIDLAQI 291
 Db 239 PDQLEAVIHPQSIHSMVDYVDSGLVLAQMGPNDRPTPIAHMAMPERFDSGVAPLDIFEV 298
 QY 292 ASLTFQKPDDEERFAPLRLARVMAARGISGAFAFNAAKEIALDHTIAGRIGLDMAAVVEE 351
 Db 299 GHMDEKPRDLKRFPCRLARVMAARGISGAFAFNAAKEIALDHTIAGRIGLDMAAVVEE 358
 QY 352 TLAVGSTDPLGKVPDALEEVLAAMDHLARRAAEE 385
 Db 359 SMAQFKPDD-----AGSLELVLAQNDQDAREVARD 387
 RESULT 6
 AAE22302
 ID AAE22302 standard; protein; 394 AA.
 AC AAE22302;
 XX 25-JUL-2002 (first entry)
 DE Methylobionas 16a sp. D-1-deoxyxylulose-5-phosphate reductoisomerase.
 XX Carotenoid; isopentenyl pyrophosphate; antheraxanthin; astaxanthin; diet;
 KM anti-oxidant; steroid; flavour; fragrance; electro-optic application;
 KM aquaculture; enzyme; D-1-deoxyxylulose-5-phosphate reductoisomerase; Dxr.
 XX
 OS Methylobionas 16a sp.
 XX
 PN WO200218617-A2.
 PD 07-MAR-2002.
 PE 04-SEP-2001; 2001WO-US27420.
 PR 01-SEP-2000; 2000US-229858P.
 PR 01-SEP-2000; 2000US-229907P.
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX Brzostowicz PC, Cheng Q, Dicosimo DJ, Koffas M, Miller ES;
 PI Odom JM, Picatagallo SK, Rouviere PE;

XX WPI: 2002-351711/38.
 DR N-PSDB: AAD35500.
 XX
 PT Producing carotenoid compounds e.g. anthraxanthin and astaxanthin, by
 PT using microorganisms having a nucleic acid molecule encoding enzymes in
 PT the carotenoid biosynthetic pathway and which metabolize single carbon
 PT substrates
 XX
 PS Claim 42; Page 112-114; 156pp; English.
 XX
 CC The invention relates to a method for producing carotenoid compounds.
 CC The method comprises a transformed metabolizing host cell, comprising
 CC suitable levels of isopentenyl pyrophosphate and a nucleic acid molecule
 CC encoding an enzyme in the carotenoid biosynthetic pathway, under the
 CC control of regulatory sequences, and contacting the host cell with carbon
 CC substrate to produce a carotenoid compound. The method is useful for
 CC producing carotenoid compounds such as anthraxanthin and astaxanthin, by
 CC using microorganism having a nucleic acid molecule encoding enzymes in
 CC the carotenoid biosynthetic pathway and which metabolize single carbon
 CC substrates. The carotenoids have potent anti-oxidant properties useful in
 CC diet, and aquaculture elements. The carotenoids are also useful as
 CC intermediates in the synthesis of steroids flavours and fragrances and
 CC compounds for potential electro-optic applications. The present sequence
 CC is Methylomonas 16a sp. D-1-deoxyxylulose-5-phosphate reductoisomerase
 CC (dxr) enzyme used in the invention.
 XX
 SQ Sequence 394 AA:
 Query Match 39.2%; Score 778; DB 23; Length 394;
 Best Local Similarity 41.6%; Pred. No. 1.7e-70;
 Matches 164; Conservative 75; Mismatches 139; Indels 16; Gaps 5;
 OY 1 MSLSIFGATGSGESTFPLVNRKGGPEAFRYVALTGCNIRRLAMARALKAELAVTAH 60
 DB 1 MGICITLGTSGTSTLDVVARH--PDKYQVVALTANGNIDALYEQCLAHHEPVAVVM 58
 OY 61 EDCPLALRELAGT--GTEVAGGAQAIADAD-RPADWTMSAIVGAGLVPGMRALKKG 116
 DB 59 ESKVAEFKORIAAPVADIKVLSGSEALQOVATLENDVTYMAIVGAGLPTLAARKG 118
 OY 117 RTLALANKESLVTAGOLMRTAENGATILPVDSHSAVQAL----AGEPTACVERV 171
 DB 119 KTVLLANKALVMSGQIFMQAVSDSGAVLLPIDSEHNALFQCPAGYTGHTAKQARRIL 178
 OY 172 ITASGPFPRDMSLERIRACTVAEAOAHPMWMSGORISDSASFKNALIEITREFEGFE 231
 DB 179 LTAASGPFPRTRPIETLSVTTPDOAVAHKPKMKRKISVSATMMNGLEIEICLLFNM 238
 OY 232 PDRIEAVNHPOSTIVHAMVFCDDGLMAHLGPDADMRHAIQFALNMPGRGEVPAVRIADLOI 291
 DB 239 PQDIEVYVHPOSTIHSMDVYDGSVLAQMGNDPMRTPPIAHAMAMPREFDSGVALDIFEV 298
 OY 292 ASLTFOKPEERPPALRLADYMAARGLSGAFAFNAAKETALDHFITAGRTIGFLDMAVVEE 351
 DB 299 GHMDFEKPLKRPCLRLAYEAIKSGGIMPTVLANAETAVEAEFLNEEVKFTDIAVIER 358
 OY 352 TLAGVSTDPFLFGKVPDALEEVLAHDLARRAAEE 385
 DB 359 SMAQFRPDD-----AGSLLELVLAQADDAAREVARD 387
 RESULT 7
 AAU80326
 ID AAU80326 standard; Protein; 394 AA.
 XX
 AC AAU80326;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Methylomonas 16a ORF2 dxr protein sequence.
 XX
 KM Isoprenoid biosynthetic enzyme; isoprenoid compound; feed additive;

KM keratenoide; pigment; flavour; fragrance; open reading frame 2; ORF2;
 KW dxr; 1-deoxyxylulose-5-phosphate reductoisomerase enzyme.
 XX
 OS Methylomonas sp.
 XX
 PN WO200220733-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 29-AUG-2001; 2001WO-US26852.
 XX
 PR 01-SEP-2000; 2000US-229907P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 PI Cheng Q, Koffas M, Norton KC, Odom JM, Picataggio SK, Rouviere PE;
 PI Schenzle A, Tomb J;
 DR WPI: 2002-383051/41.
 DR N-PSDB: ABR50082.
 XX
 PT Novel nucleic acid molecule encoding a isoprenoid biosynthetic enzyme,
 PT isolated from Methylomonas 16a, useful for the production of isoprenoid
 PT compounds -
 XX
 PS Claim 4; Page 68-70; 84pp; English.
 XX
 CC The present invention relates to a new nucleic acid molecule encoding
 CC an isoprenoid biosynthetic enzyme isolated from Methylomonas 16a.
 CC The invention is useful for obtaining a nucleic acid molecule
 CC encoding an isoprenoid compound biosynthetic enzyme, and for the
 CC microbial production of isoprenoid compounds. The molecules of the
 CC invention are also useful for regulating isoprenoid biosynthesis in an
 CC organism and for producing recombinant organisms for producing various
 CC isoprenoid compounds. The nucleic acid is also useful for feed additive,
 CC for the production of ketonoids and their derivatives, isoprenoid,
 CC intermediates, and as pure products useful as pigments, flavours and
 CC fragrances. The present amino acid sequence represents the Methylomonas
 CC 16a open reading frame 2 (ORF2) dxr (1-deoxyxylulose-5-phosphate
 CC reductoisomerase enzyme) protein of the invention, as described above.
 XX
 SQ Sequence 394 AA:
 Query Match 39.2%; Score 778; DB 23; Length 394;
 Best Local Similarity 41.6%; Pred. No. 1.7e-70;
 Matches 164; Conservative 75; Mismatches 139; Indels 16; Gaps 5;
 OY 1 MSLSIFGATGSGESTFPLVNRKGGPEAFRYVALTGCNIRRLAMARALKAELAVTAH 60
 DB 1 MGICITLGTSGTSTLDVVARH--PDKYQVVALTANGNIDALYEQCLAHHEPVAVVM 58
 OY 61 EDCPLALRELAGT--GTEVAGGAQAIADAD-RPADWTMSAIVGAGLVPGMRALKKG 116
 DB 59 ESKVAEFKORIAAPVADIKVLSGSEALQOVATLENDVTYMAIVGAGLPTLAARKG 118
 OY 117 RTLALANKESLVTAGOLMRTAENGATILPVDSHSAVQAL----AGEPTACVERV 171
 DB 119 KTVLLANKALVMSGQIFMQAVSDSGAVLLPIDSEHNALFQCPAGYTGHTAKQARRIL 178
 OY 172 ITASGPFPRDMSLERIRACTVAEAOAHPMWMSGORISDSASFKNALIEITREFEGFE 231
 DB 179 LTAASGPFPRTRPIETLSVTTPDOAVAHKPKMKRKISVSATMMNGLEIEICLLFNM 238
 OY 232 PDRIEAVNHPOSTIVHAMVFCDDGLMAHLGPDADMRHAIQFALNMPGRGEVPAVRIADLOI 291
 DB 239 PQDIEVYVHPOSTIHSMDVYDGSVLAQMGNDPMRTPPIAHAMAMPREFDSGVALDIFEV 298
 OY 292 ASLTFOKPEERPPALRLADYMAARGLSGAFAFNAAKETALDHFITAGRTIGFLDMAVVEE 351
 DB 299 GHMDFEKPLKRPCLRLAYEAIKSGGIMPTVLANAETAVEAEFLNEEVKFTDIAVIER 358
 OY 352 TLAGVSTDPFLFGKVPDALEEVLAHDLARRAAEE 385
 DB 359 SMAQFRPDD-----AGSLLELVLAQADDAAREVARD 387

Db	359	SWAOKFPDD----	AGSLELVLAQADDAEVAR	387
RESULT 8				
ID	AAO21884	standard;	Protein; 396 AA.	
AC	AAO21884			
XX	AAO21884;			
XX				
XX	13-SEP-2002	(first entry)		
DE				
XX	Isoprenoid-related	protein sequence SEQ ID NO 117.		
XX	Isoprenoid; CoQ(10); 1-deoxyxylulose-5-phosphate synthase; DXS; DDS;			
XX	decaprenyl diphosphate synthase.			
OS	Pseudomonas aeruginosa.			
XX	WO200226933-A2.			
PN	04-APR-2002.			
PD				
XX	28-SEP-2001; 2001WO-US30328.			
PF				
XX	29-SEP-2000; 2000US-236580P.			
PR				
XX	(CRGI) CARGILL INC.			
PA				
XX	Gokarn R, Jessen H, Zidwick MJ;			
PI				
XX	WPI: 2002-416480/44.			
DR				
XX	Substantially pure polypeptides having e.g.,			
FT	1-deoxyxylulose-5-phosphate synthase activity, useful for the			
PT	production of isoprenoids, especially CoQ(10).			
XX				
XX	Disclosure: Fig 32; 246pp; English.			
PS				
XX	The invention relates to methods and materials for the production of			
CC	isoprenoids. More particularly the invention provides isolated nucleic			
CC	acids, substantially pure polypeptides, host cells, and methods for			
CC	producing various isoprenoid compounds. The polypeptides are useful for			
CC	the production of isoprenoids, especially CoQ(10). Expressing the pure			
CC	polypeptides, which has 1-deoxyxylulose-5-phosphate synthase (DXS)			
CC	activity or decaprenyl diphosphate synthase (DDS) activity, is useful for			
CC	increasing production of CoQ(10) in a cell having endogenous DDS			
CC	activity. This sequence represents a protein relating to the isoprenoid			
CC	production of the invention.			
XX				
XX	Sequence 396 AA;			
XX				
XX	Query Match 39.1%; Score 775.5; DB 23; Length 396;			
XX	Best Local Similarity 42.8%; Pred. No. 3,2e-70;			
XX	Matches 166; Conservative 73; Mismatches 134; Indels 15; Gaps 5			
QY	4	LSIFGATGSGIGSTEDLVVRKKGPEAFRTVALTGRNIRRLAEMARALKAELAVTAHEDC 63		
DB	7	ISVLGATGSGIGSTLDVVGRRH--PRRYEAFALTGFSRLAELEALGCRHRRVYAVVPEQQA 64		
QY	64	LPALREALAGTG--TEVAGGAQAIAEADRP-ADWTMSAIVGAGLVPGGRALKHGRTLA 120		
DB	65	AIALGSGSLAAAIRTRVLEGEQALCEVSAPEVDYMAAIVGAAGLPSTLAAVEAKRKVL 124		
QY	121	LANKESLVYAGQLMRTAQNENCATILPVNSESAVFOALA-----GEDTACVRRVITTS 175		
DB	125	LANKALVMSGALPMQAVKRSQAVLLPIDSERNATIQSLPRNTADGLERYGVARRILLTAS 184		
QY	176	GGPEFWDMSLERIACVVAEQAHPMWSMGQRISDSASFNKALDELIEREFEGFEPDR 235		
DB	185	GGPEFETPELEPLASATYLPEDQCAHPMWSMGQRISDSASMNKGLLEIDECWLFDAQPSQV 244		
QY	236	EAVYFQSTVIAHMGFCDDGLAHNLGPRADMRALIGALVMPGSGEVPVARIDLAQIASLT 295		

Db	245	EVH1PQSVHSMVDVYVGSVIAQLGNPDMRPTPISYAAMWPERIDSGVSPIDMFAYGRLD	304
Qy	236	FOKPEEFPALRLARDVMAARGLSGAAPNAKKEIALDHF1AGTGF1DMAVVEETLAG	355
Db	305	FORDPEQFPFCIRLASQAETGSGAPPAULMANEVAFAFLERH1RFSDAVIEDVLNR	364
Qy	356	VSTDP1PFCVDPDALEEVLAMDLIARRAA	383
Db	365	EAVTAV-----ESLDQVLAADRRARSVA	387
RESULT 9			
AA	AB01369	standard; Protein: 473 AA.	
XX	AA	AB01369;	
AC	AA	AB01369;	
XX	DT	16-NOV-2000 (first entry)	
XX	DT	16-NOV-2000 (first entry)	
XX	DE	1-deoxy-D-xylulose-5-phosphate reductoisomerase sequence.	
XX	KM	1-deoxy-D-xylulose-5-phosphate reductoisomerase; transgenic plant;	
KM	KM	hepatic; isopentenyl diphosphate; mapping; breeding; phenotype;	
XX	probe.		
OS	XX	Triticum aestivum.	
PN	XX	WO200034448-A1.	
PD	XX	15-JUN-2000.	
XX	XX	03-DEC-1999; 99WO-US28616.	
XX	PR	04-DEC-1998; 98US-0110865.	
XX	PA	(DUPO) DU PONT DE NEMOURS & CO E. I.	
PI	XX	Cahoon RE, Lee J, Tao Y;	
DR	XX	WPI: 2000-431295/37.	
DR	XX	N-PSDB: AAAA47412.	
XX	PT	Novel polynucleotide encoding isopentenyl diphosphate biosynthetic	
PT	PT	enzymes useful for producing transgenic plants with altered isopentenyl	
PT	PT	diphosphate levels and for selecting polynucleotides affecting	
XX	XX	expression of the enzyme	
PS	XX	Claim 2: Page 55-57; 63pp; English.	
XX	CC	Nucleic acids encoding 1-deoxy-D-xylulose-5-phosphate	
XX	CC	reductoisomerase are useful for creating transgenic plants with	
CC	CC	altered levels of plastid isopentenyl diphosphate in cells.	
CC	CC	The 1-deoxy-D-xylulose 5-phosphate reductoisomerase polypeptides are	
CC	CC	useful as targets for identifying inhibitors of the enzyme which are	
CC	CC	useful as herbicides. Nucleic acid fragments of 1-deoxy-D-xylulose-5-	
CC	CC	phosphate reductoisomerase are useful as probes for genetically and	
CC	CC	physically mapping genes useful in plant breeding for developing	
XX	XX	lines with desired phenotypes.	
SO	Sequence	473 AA;	
Qy	Query Match	38.1%; Score 756.5; DB 21; Length 473;	
Db	Best Local Similarity	42.3%; Pred. No. 3.6e-68;	
Db	Matches 166; Conservative	69; Mismatches 138; Indels 19; Gaps	
Qy	2	RSLSIFGATGSGESTFPLVMKGGPEAFRYALTGGRNIRLAEAMARKLELAVTAHE	61
Db	76	KPISTVGTSGTGTOTDIAVE--NPDKFRVALAGSNVTLAQQVTEFKELVAVRNE	133
Qy	62	DCLPLRALALGTG--TEVAGAAQAIIEAADPP-ADWTMSAIVGAAGVPGGRALKHGRT	118
Db	134	SLNELKRALLAGCEMPELIIPEGQSVIEYARHPDAVTVYTGIVGCAGLKPYAALIEAGKD	193

CC increasing production of Coq(10) in a cell having endogenous DDS
 CC activity. This sequence represents a protein relating to the isoprenoid
 CC production of the invention.
 XX
 SQ Sequence 394 AA;
 Query Match 37.5%; Score 743.5; DB 23; Length 394;
 Best Local Similarity 43.0%; Pred. No. 5.8e-67;
 Matches 169; Conservative 58; Mismatches 151; Indels 15; Gaps 5;
 YY 4 LSTIGANGSGESTFEDLVMRKSGPEAFRTVALTGGRNIRLAEKARALKAEATVAHEDC 63
 Db 6 LTLIGSGTSGESTLVDSRHH--PEKFRVFLAGHKQVEKLAACQTFPHRYAVVADEH 63
 YY 64 LPALREALA--GTGTEVAGAGATAEADRP--ADMTSAIYGAAGLVGMRLKHGRTLA 120
 Db 64 AARLEALLKRGCTATQYLHGQALVDVASADEVSGWCAIYGAAGLVGSALAAKRGRTIY 123
 YY 121 LANKESTLVTAQGLMRTAENGATILPVDSHSAVFQALAGEDTA-----CVERVITTA 175
 Db 124 LANKESTLVSGALPMTARANGAAVLVPDSEHNAVFQVLPDYGRNLNENHIAIITLAS 183
 YY 176 GGFPRDMSLERIRACTVAEQAHPNWSGGRISIDSAMFNKALELLETREFFGFEPPDR 235
 Db 184 GGFPLTADLNTFDSITPDQAVKHPNWRKGRKITSVDSATMANKGLELLEAHKFLNCPDKL 243
 YY 236 EAVVHPQSIYHAWVGFCDGGLMAHLGPADMRHAIGFALMPRGCEVPYARIDLAQIASLT 295
 Db 244 EVVHPQSIYHAWVGFCDGGLMAHLGPADMRHAIGFALMPRGCEVPYARIDLAQIASLT 303
 YY 296 FOKPDERFPALRLARVMAARGLSGAAFNAKRIALDHFIAGRIGFLDMAAAYVEETLAG 355
 Db 304 FOKPDERFPALRLARVMAARGLSGAAFNAKRIALDHFIAGRIGFLDMAAAYVEETLAG 363
 YY 356 VSTDPLFGKVPDALEVLAMDLARRAAEEAAG 388
 Db 364 DFSDGI-----GDIGGLLAQDARTRAAARATIG 391
 RESULT 12
 AAB01362
 ID AAB01362 standard; Protein; 473 AA.
 XX
 AC AAB01362;
 XX
 DT 16-NOV-2000 (first entry)
 XX
 DE 1-deoxy-D-xylojucose-5-phosphate reductoisomerase sequence.
 XX
 KM 1-deoxy-D-xylojucose-5-phosphate reductoisomerase; transgenic plant;
 KM herbicide; isopentenyl diphosphate; mapping; breeding; phenotype;
 KM probe.
 XX
 OS *Oryza sativa*.
 XX
 PN WO200034448-A1.
 PD 15-JUN-2000.
 XX
 PF 03-DEC-1999; 99MO-US28616.
 XX
 PR 04-DEC-1998; 98US-0110865.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.
 XX
 PI Cahoon RE, Lee J, Tao Y;
 XX
 DR WPI, 2000-431295/37.
 DR N-PSDB; AAA47405.
 XX
 PT Novel polynucleotide encoding isopentenyl diphosphate biosynthetic
 PT enzymes useful for producing transgenic plants with altered isopentenyl
 PT diphosphate levels and for selecting polynucleotides affecting

PT expression of the enzyme
 XX
 PS Claim 1; Page 41-43; 63pp; English.
 XX
 CC Nucleic acids encoding 1-deoxy-D-xylojucose-5-phosphate
 CC reductoisomerase are useful for creating transgenic plants with
 CC altered levels of plastid isopentenyl diphosphate in cells.
 CC The 1-deoxy-D-xylojucose-5-phosphate reductoisomerase polypeptides are
 CC useful as targets for identifying inhibitors of the enzyme which are
 CC useful as herbicides. Nucleic acid fragments of 1-deoxy-D-xylojucose-5-
 CC phosphate reductoisomerase are useful as probes for genetically and
 CC physically mapping genes useful in plant breeding for developing
 CC lines with desired phenotypes.
 XX
 SQ Sequence 473 AA;
 Query Match 37.4%; Score 742; DB 21; Length 473;
 Best Local Similarity 41.1%; Pred. No. 1.1e-66;
 Matches 164; Conservative 74; Mismatches 139; Indels 22; Gaps 7;
 YY 2 RSLIFGATGSGESTFEDLVMRKSGPEAFRTVALTGGRNIRLAEKARALKAEATVAHE 61
 Db 77 KPSIVGSGTSGIGQTLDIYAE--NPDKFRVVALAAGSNVTLADQYKTFKPLVAVRNE 134
 YY 62 DCLPALREALAGT--TEVAGAGATAEADRP--ADMTSAIYGAAGLVGMRLKHGRT 118
 Db 135 SLVDELKEALADCDKPEEIIIPGEGVIEVARHPDAVTVVGIYGCAGLKEPTVAAIENGKD 194
 YY 119 LALANKESTLVTAQGLMRTAENGATILPVDSHSAVFQALAGEDTACVERVITTAAGGP 178
 Db 195 IALANKESTLVAGGFVPLAOKKHVKILPLADSEHSAIFQCIQGLEPALKRIITLASGA 254
 YY 179 FRDMSLERIRACTVAEQAHPNWSGGRISIDSAMFNKALELLETREFFGFEPPDRIEAV 238
 Db 255 FRDMPVDKLEKVVADALKHPNWRKGRKITSVDSATLFNKGLEVEAHYLGAEVDIEIV 314
 YY 239 VHPQSIYHAWVGFCDGGLMAHLGPADMRHAIGFALMPRG--GEVYARIDLAQIASLT 295
 Db 315 IHPOSIIHSMLETODSSVLAQLGMPDMRIPTLYTSMWPDRIYCEVTPMRLDCKIGSLT 374
 YY 296 FOKPDERFPALRLARVMAARGLSGAAFNAKRIALDHFIAGRIGFLDMAAAYVEETLAG 355
 Db 375 FKAPDNKRYPSMDLAVAGRAGGTMTGVLNANAKVALELIDEKITGLDIFKVELTICDA 434
 YY 356 ----VSTDPLFGKVPDALEVLAMDLARRAAEEAAGIR 390
 Db 435 HRNELVTRP-----SLEETIHYDLWAR---EYAAASLQ 463
 RESULT 13
 AAY97197
 ID AAY97197 standard; Protein; 399 AA.
 XX
 AC AAY97197;
 XX
 DT 04-DEC-2000 (first entry)
 XX
 DE A. thaliana partial 1-deoxy-D-xylojucose-5-phosphate reductoisomerase.
 XX
 KM 1-deoxy-D-xylojucose-5-phosphate reductoisomerase; peppermint; isoprenoid;
 KM synthesis; metabolism; chlorophyll; terpenoid; insecticidal; aroma;
 KM adhesive; ink; polymer; cytosolic; antiparasitic; plant.
 XX
 OS *Arabidopsis thaliana*.
 XX
 PN WO200046346-A1.
 PD 10-AUG-2000.
 XX
 PF 27-JAN-2000; 2000MO-US02185.
 XX
 PR 03-FEB-1999; 99US-0118349.
 XX

ID	Accession	Protein	Length	Score	DB	Length	Score	DB
DB	368	HRNELVWTSPTSP-----SLEEIVHYDLMAFREYA	392					
RESULT 15								
ID	AAB01370	standard; Protein; 406 AA.						
XX	AAB01370							
XX	16-NOV-2000	(first entry)						
XX	1-deoxy-D-xylulose-5-phosphate reductoisomerase.							
XX	1-deoxy-D-xylulose-5-phosphate reductoisomerase; transgenic plant; herbicide; isopentenyl diphosphate; mapping; breeding; phenotype; probe.							
XX	Arabidopsis thaliana.							
XX	MO200034448-A1.							
XX	15-JUN-2000.							
XX	03-DEC-1999;	99WO-US28616.						
XX	04-DEC-1998;	98US-0110865.						
XX	(DUPO) DU PONT DE NEMOURS & CO E. I.							
XX	Cahoon RE, Lee J, Tao Y;							
XX	WPI; 2000-431295/37.							
XX	Novel polynucleotide encoding isopentenyl diphosphate biosynthetic enzymes useful for producing transgenic plants with altered isopentenyl diphosphate levels and for selecting polynucleotides affecting expression of the enzyme							
XX	Example 3; Page 57-58; 63pp; English.							
XX	Nucleic acids encoding 1-deoxy-D-xylulose-5-phosphate reductoisomerase are useful for creating transgenic plants with altered levels of plastid isopentenyl diphosphate in cells.							
XX	The 1-deoxy-D-xylulose-5-phosphate reductoisomerase polypeptides are useful as targets for identifying inhibitors of the enzyme which are useful as herbicides. Nucleic acid fragments of 1-deoxy-D-xylulose-5-phosphate reductoisomerase are useful as probes for genetically and physically mapping genes useful in plant breeding for developing lines with desired phenotypes.							
XX	Sequence 406 AA;							
XX	Query Match	37.4%; Score 741.5; DB 21; Length 406;						
XX	Best Local Similarity	41.3%; Pred. No. 9.6e-67;						
XX	Matches 162; Conservative 71; Mismatches 140; Indels 19; Gaps 6;							
XX	2	RLSLFTGATGSGISTFPLVMKKGGEPAFRVYALTGGRNIRRLAEMARLKAELAVTAHE 61						
XX	10	KPISIVGSGTSGIGTQDIIAAE--NPDKFRVVALAAGSNVTLADQVRRFKPALVAVRE 67						
XX	62	DCLPALRALALGTG--TEVAGGAQAIAEAADRP-ADWIMSAIVGAAGVPGRAKAKHRT 118						
XX	68	SLINELKALALADLDYKLEIIPGEDGVIEVARHPEVAVTVTGGCAGAKPTVAALAEAKD 127						
XX	119	LALANKESILVAGOLIMLTAOENGATILPVDSHSASFQALAGEPTACVERVIITASGP 178						
XX	128	IALANKETLINGGPRVLEPLANKHNKIIIPADSEHSATIQCTIOGLPEGALRKITITASGGA 187						
XX	179	FRDMSLERIRACTAEAOAHBPMWSMGORISIDSAMFNKALELITREFFGFEEDPRIAY 238						
XX	188	FRDPMVEKLEKVAADAKLHPNMWNGKITVDSATLFFKGLVIEAHYLFAGYEYDIEIV 247						
XX	239	VHPOSIYAHMAGFCDDGSLMAHLGPRDMHAIIGFALNMFGR---GEVPARIDLAOIASLT* 295						

Db 248 IHPOIISHMLETQDSSVLAQIGMPDMRPLPLYTMSMDRNPCESEVYWPRLDCKIGSLT 307

OY 236 FQKPEEEFPPLRLRLARDYMAARGISGAFAFNAKTALDHTIAGTIGFLDMAAYVEETL-- 353

Db 308 FKPPNNVYVPSMDLAIYAAGRAGGTMTGVLSTANKEAAVEFMIDEKISYLDLIFKVELTCDK 367

OY 354 --AGVSTDPGLFGKVPDALEEVLTADHDLARRAA 383

Db 368 HRLNELVTSF-----SLEELVHIVDLMAREYA 392

```
Search completed: April 15, 2003, 13:57:45
Job time : 39 secs
```

```

Db      248  IHPOSIIHSMIETQDSSVLAQLGWPDMRRLPILYTMSPDRVPCSEVTVWPRLDLCKLGSLT 307
Oy      296  FQKPEEERFPALRLARDYMAARGLSGAAFNNAKEIALDHFIAGRIGFLDMAAYVEETL-- 353
Db      308  FKPPDNVKKYPPSMDLAYAAGRAGGTMTGVLSPAANEKAVEMFIDEXISYLDLfkvVELTCDK 367
Oy      354  --AGVSTDPLFGKVPDALLEVLAMDHLARRAA 383
Db      368  HRNELVTSP-----SLEELVHYDLMAREYA 392

Search completed: April 15, 2003, 13:57:45
Job time : 39 secs

```

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2003, 13:56:29 ; Search time 21 Seconds

(without alignments)
1803.665 Million cell updates/sec

Title: US-09-673-198-30

Perfect score: 1984

Sequence: 1 MRSLSIFGATGSGESTFDL.....MDHLARRAAEAAGLRQKR 394

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: PIR_73:**

1: pir1:**

2: pir2:**

3: pir3:**

4: pir4:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	994	50.1	397	2 E97672	1-deoxy-D-xylose
2	994	50.1	397	2 AC2897	1-deoxy-D-xylose
3	899.5	45.3	399	2 H87486	1-deoxy-D-xylose
4	775.5	39.1	396	2 E83188	1-deoxy-D-xylose
5	750.5	37.8	394	2 A81229	1-deoxy-D-xylose
6	743.5	37.5	394	2 B82000	1-deoxy-D-xylose
7	741.5	37.4	406	2 T52570	1-deoxy-D-xylose
8	740	37.3	396	2 H82728	1-deoxy-D-xylose
9	733	36.9	398	2 G90650	1-deoxy-D-xylose
10	733	36.9	398	2 G85501	1-deoxy-D-xylose
11	732	36.9	398	2 E64741	1-deoxy-D-xylose
12	725	36.5	398	2 AG0128	1-deoxy-D-xylose
13	717	36.1	397	2 A64014	1-deoxy-D-xylose
14	708	35.7	398	2 AF0529	1-deoxy-D-xylose
15	707	35.6	399	2 AH2349	1-deoxy-D-xylose
16	702.5	35.4	380	2 AE1239	1-deoxy-D-xylose
17	695.5	35.0	394	2 S76331	1-deoxy-D-xylose
18	693.5	35.0	385	2 E97121	1-deoxy-D-xylose
19	682	34.4	402	2 D82099	1-deoxy-D-xylose
20	673	33.9	380	2 A11601	1-deoxy-D-xylose
21	670	33.8	388	2 B69881	1-deoxy-D-xylose
22	640	32.3	396	2 F75388	1-deoxy-D-xylose
23	631	31.8	393	2 C71304	1-deoxy-D-xylose
24	628	31.7	436	2 A70923	1-deoxy-D-xylose
25	617	31.1	398	2 F84957	1-deoxy-D-xylose
26	611.5	30.8	379	2 G81712	1-deoxy-D-xylose
27	608.5	30.7	379	2 A71562	1-deoxy-D-xylose
28	598.5	30.2	379	2 C81578	1-deoxy-D-xylose
29	596.5	30.1	379	2 G86533	1-deoxy-D-xylose

30	596.5	30.1	379	2 B72091	1-deoxy-D-xylose
31	595.5	30.0	365	2 E83952	1-deoxy-D-xylose
32	588	29.6	406	2 A87107	1-deoxy-D-xylose
33	565	28.5	380	2 F70336	1-deoxy-D-xylose
34	550.5	27.7	376	2 B72321	1-deoxy-D-xylose
35	488	24.6	368	2 H64546	1-deoxy-D-xylose
36	477	24.0	366	2 G71961	1-deoxy-D-xylose
37	468	23.6	356	2 G81278	1-deoxy-D-xylose
38	468	23.6	412	2 B87586	1-deoxy-D-xylose
39	468	23.6	412	2 H75271	1-deoxy-D-xylose
40	468	23.6	412	2 B98241	1-deoxy-D-xylose
41	468	23.6	412	2 AB3045	1-deoxy-D-xylose
42	468	23.6	412	2 B70877	1-deoxy-D-xylose
43	468	23.6	412	2 E87621	1-deoxy-D-xylose
44	468	23.6	412	2 H83018	1-deoxy-D-xylose
45	468	23.6	412	2 AE3550	1-deoxy-D-xylose

ALIGNMENTS

RESULT 1

E97672

1-deoxy-D-xylose 5-phosphate reductoisomerase (dxp reductoisomerase) [Imported] - A

C:Species: Agrobacterium tumefaciens

C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002

C:Accession: E97672

R:Gooner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Gold

A.: Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Martel,

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium

A:Reference number: A97359; PMID:11743194

A:Accession: E97672

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-397 <RUR>

A:Cross-references: GB:AE007869; PIDN:AAK8334.1; PID:G15157812; GSPDB:GND0169

C:Genetics:

A:Gene: AGR C.4736

A:Map position: circular chromosome

C:Superfamily: conserved hypothetical protein HI0807

Query Match	50.1%	Score 994	DB 2	Length 397
Best Local Similarity	50.4%	Pred. No. 1.2e-63		
Matches 197	Conservative 64	Mismatches 126	Indels 4	Gaps 1
QY	2	RSLSIFATGSGISGFTFDLVMKRGPRFAFTVALTGRNTRLAEMARALKAEIATVAHE	61	
DB	9	RKLITLIGSTGISTNTLDVVRQJGGRGFEIMALTGAGNIALLAEOARREGAQLAVTADD	68	
QY	62	DCUPALREALAGTGEVAGGAGQAIAEAADRPADMTGAIAGAGLVGMRALKHGFAL	121	
DB	69	DKYALKSALAGTGIKRAAGAGLEEAASMDAGVMAAIAIGTGLATVLAARAGADIAL	128	
QY	122	ANKESIVTAQQLMRTAENGATILPVDSHSAVFOALAGEDNACVETVITASGGEFRD	161	
DB	129	ANKKCLVAGSAGVFLRTVQKGGRIIPVDSHSAIFQGLTEYQAVRIVLTASGGEFR	168	
QY	182	WSLEIRIACIVAEQAHPNMSMGORISIDASMNKALELIERREFGFEPRDREAVHP	241	
DB	189	WSRDEMSNVTAADIAFARHPNMSMGLKVSIGSASMNKLEMEIAYLFDLPDQDVIVHP	248	
QY	242	QSIYHAWGFCGGIMAHLPADMRHIGFALNMPGGEVPAVIDLAOLASLTFOKPDE	301	
DB	249	QSIHSHVGTGDSYIAQLQSPMDRTAISTALYPERGNISVERLDPAKLARLDEAPDE	308	
QY	302	EREPALRLADYVMAARGLSCAAFAAEIALDHFIAIRIGFLDMAAVVEETLAGVSTPDL	361	
DB	309	ARFALALAMALERGLQGAALNAAEETAFHFAVAGIGTFLDMAELVEYVMDRMHD---	365	
QY	362	FGKVPDALEEVYLAADHLARRAAEAAGLRQK	392	
DB	366	GRTAETMDVFSADDEARRRALELLIATKER	395	

A:Molecule type: mRNA
A:Residues: 1-406 <SCH>
A:Cross-references: EMBL:AJ242588; NID:9486306; PIDN:CA84344.1; PID:9486307
A:Experimental source: cultivar Columbia
C:Genetics:
A:Gene: dxt
C:Superfamily: conserved hypothetical protein H10807

Query Match 37.4%; Score 741.5; DB 2; Length 406;
Best Local Similarity 41.3%; Pred. No. 1.4e-45;
Matches 162; Conservative 71; Mismatches 140; Indels 19; Gaps 6;

QY 2 RSLISFGATSGISGTFDLYMRKGGPEAFRTVALTGGRNIRRLAEMARALKAELAVTAH 61
D 10 KRISVSGTSGISGTFDLYMRKGGPEAFRTVALTGGRNIRRLAEMARALKAELAVTAH 67
QY 62 DCLPALREAL--AGTGVAGGAOIAEADRP--ADMTMSAIVGAGVPGKRALKHRT 118
D 68 SLINELKALADLYKKEIIFGEGVIEVARHPEAVYVGTIGVCGAGLKPVAALAEAGKD 127
QY 119 LALANKESLVTAQGLLMRTAENGATILPVDESHSAVFQALAGEDTACVERVITASGP 178
D 128 IALANKETLACGFEVLEPLANKHNYKILPADSEHSAIFQCIQGLPEGALRKILITASGA 187
QY 179 FRDMSLERIRACTVAEAOAHPNMWSGQRISIDSASMFNKALELITREFFGPEPRIEA 238
D 188 FRDMPVEKLEKVAADALKHPNMNKKRITVDSATLFFNKGLEVEIAHYLFGAEXDDIIV 247
QY 239 VHPDSIVAHAVGFCDDGLMAHLGPAIDMRHAIIGFALNMPGR--GEVPAIDILAOIASLT 295
D 248 IHPSIIHSMLEIQTDSVLAQGLMDMLPLITMSWPDRCSEVTMPRIDCLSLT 307
QY 296 FOKDEEFEPRLRLARDVMAARGLSGAFAFNAAKETALDHFITAGRIGLDMAAVEETL-- 353
D 308 FKRPDMKYPMDLATAAGRAGTGTGVLASANEKAVEMFIDEKISYDIFKVVETLTDCK 367
QY 354 --AGVSTDPLEFGVDALEEVTLAMDHLARRAA 383
D 368 HRNELVTSF-----SLEIVHYDLMAREYA 392

RESULT 8

H82728
1-deoxy-D-xylose 5-phosphate reductoisomerase Xf1048 [imported] - Xylella fastidiosa
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C:Accession: H82728
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: H82728
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-396 <SIM>
A:Cross-references: GB:AE003942; GB:AE003849; NID:99105990; PIDN:AAF83858.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simpon, A.J.G.; Reinach, E.L.; Artuda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Brites, M.R.S.; Bueno, M.R.P.; Camargo, A.; Camargo, L.E.A.; Carraro, D.M.; Carre, H
as-Melo, E.; Docena, C.; El-Dor, H.; Facinani, A.P.; Ferreira, A.J.S.
submitted to Genbank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Jungueira, M.L.; Kemper, E.L.; Kitzima, J.P.; Krieger, J.E.; Kurame, E.E.; Laig
chao, D.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, B
A:Authors: Martins, E.M.F.; Matsushima, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Sattelli, R.V.; Savasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshano, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:

A:Gene: Xf1048
C:Superfamily: conserved hypothetical protein H10807

Query Match 37.3%; Score 740; DB 2; Length 396;
Best Local Similarity 42.4%; Pred. No. 1.7e-45;
Matches 165; Conservative 60; Mismatches 154; Indels 10; Gaps 5;

QY 1 MRSLSIFGATSGISGTFDLYMRKGGPEAFRTVALTGGRNIRRLAEMARALKAELAVTAH 60
D 5 IRNVAIVGATSGISGALADVLARH--PROFHVSLLAQGRDALLALCHTRPDHVAIGD 62
QY 61 EDCLPALREAL--AGTGVAGGAOIAE--AADRADMTMSAIVGAGVPGKRALKHGR 117
D 63 ATLYTTLGDLNAGLAKAVAGEAALAEVASTCDDVVAIVGAGVPGKRALKHGR 122
QY 118 TLALANKESLVTAQGLLMRTAENGATILPVDESHSAVFQALAGEDTACVERVITASGP 177
D 123 RLLANKESLVTAQGLLMRTAENGATILPVDESHSAVFQALAGEDTACVERVITASGP 182
QY 178 FRDMSLERIRACTVAEAOAHPNMWSGQRISIDSASMFNKALELITREFFGPEPRIEA 237
D 183 FRDMPVEKLEKVAADALKHPNMNKKRITVDSATLFFNKGLEVEIAHYLFGAEXDDIIV 242
QY 238 VHPDSIVAHAVGFCDDGLMAHLGPAIDMRHAIIGFALNMPGR--GEVPAIDILAOIASLT 297
D 243 LVHPSIIHSMLEIQTDSVLAQGLMDMLPLITMSWPDRCSEVTMPRIDCLSLT 302
QY 298 FKDEEFEPRLRLARDVMAARGLSGAFAFNAAKETALDHFITAGRIGLDMAAVEETL 357
D 303 RPDTEFSCRLRLARDVMAARGLSGAFAFNAAKETALDHFITAGRIGLDMAAVEETL 359
QY 358 TDPLFGKVPDALEEVTLAMDHLARRAAEA 386
D 360 TLPRY--EADTLETLITVDTERTITTHAA 386

RESULT 9

G90650
1-deoxy-D-xylose 5-phosphate reductoisomerase [imported] - Escherichia coli (strain
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: G90650
R:Havashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gawara, N.; Yasunaga, T.; Kohara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: G90650
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-398 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA833598.1; PID:913359631; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: Ecs0175
C:Superfamily: conserved hypothetical protein H10807

Query Match 36.9%; Score 733; DB 2; Length 398;
Best Local Similarity 40.7%; Pred. No. 5.5e-45;
Matches 162; Conservative 72; Mismatches 144; Indels 20; Gaps 5;

QY 1 MRSLSIFGATSGISGTFDLYMRKGGPEAFRTVALTGGRNIRRLAEMARALKAELAVTAH 60
D 1 MKQTLITGSGISGTFDLYMRKGGPEAFRTVALTGGRNIRRLAEMARALKAELAVTAH 58
QY 61 EDCLPALREAL--AGTGVAGGAOIAE--AADRADMTMSAIVGAGVPGKRALKHGR 117
D 59 EASKKLTKMLQGGKSTEVLSGQACDMALEEDVDVMAIVGAGVPGKRALKHGR 118
QY 118 TLALANKESLVTAQGLLMRTAENGATILPVDESHSAVFQALAGEDTACV 167
D 119 TILANKESLVTCGLRFDADAKQSKAQLLPDSEHNAIFQSLPPQICHNGLYADLEONGV 178

C:Superfamily: conserved hypothetical protein HI0807
C:Keywords: oxidoreductase

Query Match 36.5%; Score 725; DB 2; Length 398;
Best Local Similarity 40.7%; Pred. No. 2,1e-44;
Matches 162; Conservative 69; Mismatches 147; Indels 20; Gaps 5;

```
QY 1 MRSLSIFGATSGISESTFDLVNRRKGPEAFRTVALTGGRNIRLAEKARLAEVATYAH 60
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MKQTLITGSGTSGISNSTLSVY--RANPEFKYALVAGNRVREMAQOCLEFSRYAAMD 58
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 EDCPLALREALA--GTGEVAGGAQALAE--AADRPADMTSAIYAGAGLVGMRLAKGR 117
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 59 EHSKAKSLRLAEQSGSDTEVSGETAACELALDDVDQMAIYAGIAGLPSTLAIRAGK 118
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 118 TLALANKESLYTAGQLMRTAQENGATILPVDESHSAVFQAL-----AGEDTACY 167
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 QVLANKESSLITGCKLPFDEVKRSRAQLPIDSERHNAIFQSLPERIQKQLGYSLSLNGV 178
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 168 ERVITTAGSGPFRDWSLERIRACTVAEQAQHPNMSGCRISIDSASFENKALEITREF 227
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 SRILITGSGPFRPTLSQFSDVTPDQCAHPNMSGKRISVDSATVMNKGLEYIEARWL 238
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 228 FGFEPRDRLEAVVHQSIYHNAVVGCDGGLMAHLGPADMRHAIIGFALNMPGGEVPAVARID 287
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 239 FNASAEQLEVLVHPQSVIHSVRYHDSITLQMGTPDMRTTIAHMAIYPMRVSSGVAFLD 298
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 288 LAQIASLTFQKPDDEERPALRLARDVMAARGLSGAFAFNAKEIALDHFIAGRIGFLDMAA 347
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 299 FCKVAGLTFTPDYQRYPCKLALDADACNAGQAATTALANAEISVMAFLDSKIRTDIEV 358
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 348 VVEETLAGVSTDPPLFGKVPDALVEVLAMDHLARRAAE 385
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 359 INRTVEG---LTLSEPTVSEEVLYIDRRARVAAQ 391
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
    RESULT 13
    A64014
    conserved hypothetical protein HI0807 - Haemophilus influenzae (strain Rd KW20).
    C:Species: Haemophilus influenzae
    C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 29-Sep-1999
    C:Accession: A64014
    R:Flieschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirnness, E.F.; Kerlavage, A.;
    Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.;
    D.M.; Brandon, R.C.; Fine, L.D.; Fitchman, J.L.; Fullmann, J.L.; Geoghagen, N.S.M.
    Science 269, 496-512, 1995
    A:Authors: Guehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
    A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
    A:Reference number: A64000; MUID:95350630; PMID:7542800
    A:Accession: A64014
    A:Status: nucleic acid sequence not shown; translation not shown
    A:Molecule type: DNA
    A:Residues: 1-397 <TRIG>
    A:Cross-references: GB:U32763; GB:I42023; NID:g1573817; PIDN:AAC22466.1; PID:g1573819; T
    C:Superfamily: conserved hypothetical protein HI0807
```

Query Match 36.1%; Score 717; DB 2; Length 397;
Best Local Similarity 40.2%; Pred. No. 7.7e-44;
Matches 159; Conservative 72; Mismatches 145; Indels 20; Gaps 5;

```
QY 2 RSLISFGATSGISESTFDLVNRRKGPEAFRTVALTGGRNIRLAEKARLAEVATYAH 61
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 QNITLITGSGTSGISNSTLSVY--ENNPQKYHAFALVGGKNAVMEQCIFRPHFAALDDV 61
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 DCLPALREALAG--TGEVAGGAQALAEADR--ADMTMSAIVGAGLVGMRLAKHGR 118
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 NAAKILKEKLAHRIETEVLAGRAICELAAHPADQIMASTVAGAGLPLTSAVKAKGR 121
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 119 LALANKESLYTAGQLMRTAQENGATILPVDESHSAVFQALAGE-----DTACVE 168
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 VILANKESLYTAGQLMRTAQENGATILPVDESHSAVFQALAGE-----DTACVE 168
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 169 RVITTAGSGPFRDWSLERIRACTVAEQAQHPNMSGCRISIDSASFENKALEITREF 228
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
    RESULT 14
    A64014
    conserved hypothetical protein HI0807 - Haemophilus influenzae (strain Rd KW20).
    C:Species: Haemophilus influenzae
    C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 29-Sep-1999
    C:Accession: A64014
    R:Flieschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirnness, E.F.; Kerlavage, A.;
    Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.;
    D.M.; Brandon, R.C.; Fine, L.D.; Fitchman, J.L.; Fullmann, J.L.; Geoghagen, N.S.M.
    Science 269, 496-512, 1995
    A:Authors: Guehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
    A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
    A:Reference number: A64000; MUID:95350630; PMID:7542800
    A:Accession: A64014
    A:Status: nucleic acid sequence not shown; translation not shown
    A:Molecule type: DNA
    A:Residues: 1-397 <TRIG>
    A:Cross-references: GB:U32763; GB:I42023; NID:g1573817; PIDN:AAC22466.1; PID:g1573819; T
    C:Superfamily: conserved hypothetical protein HI0807
```

Query Match 35.7%; Score 708; DB 2; Length 398;
Best Local Similarity 40.2%; Pred. No. 3.4e-43;
Matches 160; Conservative 72; Mismatches 146; Indels 20; Gaps 6;

```
QY 1 MRSLSIFGATSGISESTFDLVNRRKGPEAFRTVALTGGRNIRLAEKARLAEVATYAH 60
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MKQTLITGSGTSGISNSTLSVY--RANPEFKYALVAGNRVREMAQOCLEFSRYAAMD 58
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 EDCPLALREALA--GTGEVAGGAQALAE--AADRPADMTSAIYAGAGLVGMRLAKGR 117
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 59 EHSKAKSLRLAEQSGSDTEVSGETAACELALDDVDQMAIYAGIAGLPSTLAIRAGK 118
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 118 TLALANKESLYTAGQLMRTAQENGATILPVDESHSAVFQAL-----AGEDTACY 167
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 QVLANKESSLITGCKLPFDEVKRSRAQLPIDSERHNAIFQSLPERIQKQLGYSLSLNGV 178
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 168 ERVITTAGSGPFRDWSLERIRACTVAEQAQHPNMSGCRISIDSASFENKALEITREF 227
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 SRILITGSGPFRPTLSQFSDVTPDQCAHPNMSGKRISVDSATVMNKGLEYIEARWL 238
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 228 FGFEPRDRLEAVVHQSIYHNAVVGCDGGLMAHLGPADMRHAIIGFALNMPGGEVPAVARID 287
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 239 FNASAEQLEVLVHPQSVIHSVRYHDSITLQMGTPDMRTTIAHMAIYPMRVSSGVAFLD 298
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 288 LAQIASLTFQKPDDEERPALRLARDVMAARGLSGAFAFNAKEIALDHFIAGRIGFLDMAA 347
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 299 FCKVAGLTFTPDYQRYPCKLALDADACNAGQAATTALANAEISVMAFLDSKIRTDIEV 358
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 348 VVEETLAGVSTDPPLFGKVPDALVEVLAMDHLARRAAE 385
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 359 L---NLAVLERMDL--HEPASVDVDTLOYDALAREVARK 391
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
    RESULT 14
    AF0529
    1-deoxy-D-xylulose 5-phosphate reductoisomerase (EC 1.1.1.-) [imported] - Salmonella
    C:Species: Salmonella enterica subsp. enterica serovar Typh
    A:Note: this species has also been called Salmonella typh
    C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
    C:Accession: AF0529
    R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
    th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
    , S.; Moule, S.; O'Gaora, P.
    Nature 413, 848-852, 2001
    A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
    A.; Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
    A:Reference number: AB0502; PMID:11677608
    A:Accession: AF0529
    A:Status: preliminary
    A:Molecule type: DNA
    A:Residues: 1-398 <PAR>
    A:Cross-references: GB:AL513382; PIDN:CAD08678.1; PID:g16501501; GSPDB:GN00176
    C:Genetics:
    A:Gene: dxr
    C:Superfamily: conserved hypothetical protein HI0807
    C:Keywords: oxidoreductase
```

Query Match 35.7%; Score 708; DB 2; Length 398;
Best Local Similarity 40.2%; Pred. No. 3.4e-43;
Matches 160; Conservative 72; Mismatches 146; Indels 20; Gaps 6;

```
QY 1 MRSLSIFGATSGISESTFDLVNRRKGPEAFRTVALTGGRNIRLAEKARLAEVATYAH 60
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MKQTLITGSGTSGISNSTLSVY--RANPEFKYALVAGNRVREMAQOCLEFSRYAAMD 58
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 EDCPLALREALA--GTGEVAGGAQALAE--AADRPADMTSAIYAGAGLVGMRLAKGR 117
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 59 EHSKAKSLRLAEQSGSDTEVSGETAACELALDDVDQMAIYAGIAGLPSTLAIRAGK 118
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 118 TLALANKESLYTAGQLMRTAQENGATILPVDESHSAVFQAL-----AGEDTACY 167
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 QVLANKESSLITGCKLPFDEVKRSRAQLPIDSERHNAIFQSLPERIQKQLGYSLSLNGV 178
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 168 ERVITTAGSGPFRDWSLERIRACTVAEQAQHPNMSGCRISIDSASFENKALEITREF 227
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 SRILITGSGPFRPTLSQFSDVTPDQCAHPNMSGKRISVDSATVMNKGLEYIEARWL 238
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 228 FGFEPRDRLEAVVHQSIYHNAVVGCDGGLMAHLGPADMRHAIIGFALNMPGGEVPAVARID 287
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 239 FNASAEQLEVLVHPQSVIHSVRYHDSITLQMGTPDMRTTIAHMAIYPMRVSSGVAFLD 298
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 288 LAQIASLTFQKPDDEERPALRLARDVMAARGLSGAFAFNAKEIALDHFIAGRIGFLDMAA 347
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 299 FCKVAGLTFTPDYQRYPCKLALDADACNAGQAATTALANAEISVMAFLDSKIRTDIEV 358
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 348 VVEETLAGVSTDPPLFGKVPDALVEVLAMDHLARRAAE 385
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 359 L---NLAVLERMDL--HEPASVDVDTLOYDALAREVARK 391
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
    RESULT 14
    AF0529
    1-deoxy-D-xylulose 5-phosphate reductoisomerase (EC 1.1.1.-) [imported] - Salmonella
    C:Species: Salmonella enterica subsp. enterica serovar Typh
    A:Note: this species has also been called Salmonella typh
    C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
    C:Accession: AF0529
    R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
    th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
    , S.; Moule, S.; O'Gaora, P.
    Nature 413, 848-852, 2001
    A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
    A.; Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
    A:Reference number: AB0502; PMID:11677608
    A:Accession: AF0529
    A:Status: preliminary
    A:Molecule type: DNA
    A:Residues: 1-398 <PAR>
    A:Cross-references: GB:AL513382; PIDN:CAD08678.1; PID:g16501501; GSPDB:GN00176
    C:Genetics:
    A:Gene: dxr
    C:Superfamily: conserved hypothetical protein HI0807
    C:Keywords: oxidoreductase
```

```

RESULT 15
AH2349
deoxyxylulose 5-phosphate reductoisomerase [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AH2349
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kunitz, T.; Sasamoto, S.; Watanabe, A.; Iriuchihara, S.;
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH2349
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1399 <KUR>
A:Cross-References: GB:BA000019; PIN:BAW/6050.1; PID:g17133487; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr4351
C:Superfamily: conserved hypothetical protein HI0807

Query Match          35.6%; Score 707; DB 2; Length 399;
Best Local Similarity 38.0%; Pred. No. 4e-43;
Matches 146; Conservative 88; Mismatches 142; Indels 8; Gaps 4;

QY 1 MRSLSTFGATGSGTGSTGEPDLMVRKGPPEAFRTALTGGRIRRLAEARALKAEIAVTAH 60
Db 2 VKSITLVGSGTGTGTLDTVSQY--PDQFRIVGLAGSNVEMLAQEIROFRPOIAISA 59
QY 61 EDCLPALREALAGTGT--VAGGAQAIAEAAD--RPADWTMSAIVGAAGLVPGRALKHGR 117
Db 60 AEKLPALQAAIKDLPQPIILGGEAGVIEVARGDAAETVTVTGCGAGLPTIAIIEAGK 119
QY 118 TLALANKESLYTGAQLMRTAENGATILPVSEHSNAVFOALAGEDTACVRYITRASG 177
Db 120 DIALANKETLIAGPVVLPVLEKRGVLLPDADESHSAIFQCIQGVPRKGLKILLTASG 179
QY 178 PFRDMSLERIRACTVAEAQAHPMWSMGORISIDSASMFNKALELIETREFGFEPPRIEA 237
Db 180 AFRDMDVERIAETVSDALKHPMSMGKRTTVSATILMKGLEIVENHFLFGLDYQDIET 239
QY 238 VVHPOSIVHNMVGFCDGGLMAHLGPADMRHAIGFALNMPGREGVPAVARIDLAQIASLT 297
Db 240 VIHPOSIIHSLIELQDTSVLAQLGMPDMRLPLLYALSMPEIRIYTDWERNLVKAGNLTF 299
QY 298 KPDEERFPALRLARDVMAARGLSGAAPNAKETAALDHFIAGRIGFLDMAAVETLAGVS 357
Db 300 EPDHQKYPGQQLAYAAAGRAGGSPAVLNANAEQVALFDEKIKFLDIPRCIELVCDRHQ 359
QY 358 TDPLFGKVPDALEVLAMHLARR 381
Db 360 ND---NCANPSLDDIILAADQMARQ 380

```

Search completed: April 15, 2003, 13:58:13
 Job time : 23 secs

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2003, 13:56:25 ; Search time 25 Seconds

(without alignments)
653.667 Million cell updates/sec

Title: US-09-673-198-30

Perfect score: 1984

Sequence: 1 MRSLSIFGATSGISGFIDL.....MDHLARRAAEAAGLRQOKR 394

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1024	51.6	391	1 DXR_RHIME	0921p6 rhizobium m
2	994	50.1	397	1 DXR_AGR75	08uc86 agrobacteri
3	899.5	45.3	399	1 DXR_CAVCR	09a709 caulobacter
4	875	44.1	388	1 DXR_ZYMO	09x7f2 zymomonas m
5	797.5	40.2	393	1 DXR_PALSO	08x2f5 ralsstonia s
6	775.5	39.1	396	1 DXR_PSEAE	09x6p6 pseudomonas
7	750.5	37.8	394	1 DXR_NEIMA	09j7g8 neisseria m
8	743.5	37.5	394	1 DXR_NEIMA	09j7g8 neisseria m
9	741.5	37.4	477	1 DXR_ARATH	09x7s9 arabidopsis
10	740	37.3	396	1 DXR_XYLFA	09pe10 xyella fas
11	733	36.9	398	1 DXR_ECO57	08x8y1 escherichia
12	732	36.9	398	1 DXR_ECOLI	08x8y1 escherichia
13	725	36.5	398	1 DXR_YERPE	08x8y1 escherichia
14	717	36.1	397	1 DXR_HAEN	08x8y1 escherichia
15	716	36.1	405	1 DXR_PASWU	08x8y1 escherichia
16	715	36.0	398	1 DXR_SALTY	08x8y1 escherichia
17	708	35.7	398	1 DXR_SALTY	08x8y1 escherichia
18	707	35.6	399	1 DXR_ANASP	08x8y1 escherichia
19	702.5	35.4	380	1 DXR_LISMO	08x8y1 escherichia
20	695.5	35.1	394	1 DXR_SYNY3	08x8y1 escherichia
21	693.5	35.0	385	1 DXR_CIOAB	08x8y1 escherichia
22	685	34.5	402	1 DXR_SYNYE	08x8y1 escherichia
23	684.5	34.5	401	1 DXR_STRCO	08x8y1 escherichia
24	683.5	34.4	402	1 DXR_VIBCH	08x8y1 escherichia
25	682	34.4	402	1 DXR_VIBCH	08x8y1 escherichia
26	673	33.9	380	1 DXR_LISIN	08x8y1 escherichia
27	671.5	33.8	384	1 DXR_CLOPE	08x8y1 escherichia
28	670	33.8	388	1 DXR_BACSU	08x8y1 escherichia
29	654.5	33.0	475	1 DXR_MENPI	08x8y1 escherichia
30	643.5	32.4	382	1 DXR_BACHD	08x8y1 escherichia
31	640	32.3	396	1 DXR_DEIRA	08x8y1 escherichia
32	631	31.8	376	1 DXR_TREPA	08x8y1 escherichia
33	628	31.7	413	1 DXR_MYCTU	08x8y1 escherichia

34	617	31.1	398	1 DXR_BUCAI	P57329 buchnera ap
35	611.5	30.8	379	1 DXR_CHLMU	09x7v8 chlamydia m
36	608.5	30.7	379	1 DXR_CHLTR	084074 chlamydia t
37	596.5	30.1	379	1 DXR_CHTRN	09x7v8 chlamydia p
38	588	29.6	406	1 DXR_MYCLE	09x7v8 chlamydia p
39	568	28.6	390	1 DXR_FUSNN	09x7v8 chlamydia p
40	565	28.5	380	1 DXR_AQUAE	066722 aquifex aeo
41	550.5	27.7	376	1 DXR_THEMA	09x7v8 chlamydia p
42	488	24.6	368	1 DXR_HELPY	P56139 helicobacte
43	477	24.0	368	1 DXR_HELPY	09x7v8 chlamydia p
44	468	23.6	356	1 DXR_CAMCE	09x7v8 chlamydia p
45	140.5	7.1	2564	1 SPCO_HUMAN	09x7v8 chlamydia p

ALIGNMENTS

RESULT 1	ID	DXR_RHIME	STANDARD	PRT	391 AA
AC	0921p6				
DT	15-JUN-2002 (Rel. 41, Created)				
DT	15-JUN-2002 (Rel. 41, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	1-deoxy-D-xylulose 5-phosphate reductoisomerase (EC 1.1.1.267) (DXP reductoisomerase) (1-deoxyxylulose-5-phosphate reductoisomerase).				
GN	DXP OR R02988 OR SMC03105.				
OS	Rhizobium meliloti (Sinorhizobium meliloti).				
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;				
OC	Rhizobiaceae; Sinorhizobium.				
OX	NCBI_TaxID=382;				
ON	(1)				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=1021;				
RX	MEDLINE=21396507; PubMed=11481430;				
RA	Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,				
RA	Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,				
RA	Godrie T., Goffeau A., Kahn D., Kils E., Lelaure V., Masny D.,				
RA	Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,				
RA	Renard C., Thebaud P., Vandenbol M., Weidner S., Gallbert F.,				
RT	"Analysis of the chromosome sequence of the legume symbiont				
RT	Sinorhizobium meliloti strain 1021."				
RL	Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).				
CC	-1- FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction				
CC	of 1-deoxy-D-xylulose-5-phosphate (DXP) to 2-C-methyl-D-erythritol				
CC	4-phosphate (MEP) (by similarity).				
CC	-1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADP(+) ->				
CC	1-deoxy-D-xylulose 5-phosphate + NADPH.				
CC	-1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second				
CC	step.				
CC	-1- SIMILARITY: BELONGS TO THE DXR FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation-				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL: AL591792; CAC47567.1; -				
DR	InterPro: IPR003821; DXP_reductoisomase.				
DR	Pfam: PF02670; DXP_reductoisom; 1.				
KW	TIGRPFAM: TIGR00243; Dxr; 1.				
KW	ISOPRENE BIOSYNTHESIS; Oxidoreductase; NADP; Complete proteome.				
FT	NE_BIND 14 21 NADPH (POTENTIAL).				
SQ	SEQUENCE 391 AA: 41741 MW: 280360 CDF026757D CRC64;				
QY	Query Match 51.6%; Score 1024; DB 1; Length 391;				
QY	Best Local Similarity 53.9%; Pred. No. 4.2e-65;				
QY	Matches 207; Conservative 60; Mismatches 113; Indels 4; Gaps 1;				
QY	2 RSLSIFGATSGISGFIDLVMRKGPGEAFRTVALTGGNRIRRLAEMARALKELAVTAHE 61				

```

Db 9 RRLTILSTGSGTSTLDVIERLGRDFEIALTGNNINILAEQARRIGAEIATVAD 68
QY 62 DCLPALREALAGTGEVAGGAOATAEADRPADMTMSAIVAGACIVPMRLKRGRTAL 121
Db 69 DRYELKDAISGSGLEVAAGRSGLTEAERDAGWMAIVAGGFLAARRKADIAL 128
QY 122 ANKESLVATAGOLMRTAENGATILPVDESHSAFVALAGEDTACVERVITTAAGGPPRD 181
Db 129 ANKCVISAGSLFIDVAEAGGRLIPVDESHALFOVLENOGRHVAEIVITLASSGPPRT 188
QY 182 WSLERIRACTYAEAOAHPMNSMGORISIDSAMFNKALELIEFEFGFEPDRIEAVVHP 241
Db 189 KTLDEMRTVADVAPARHPNMSGKISIDSAMFNKALEMIEARHLFRLEPQIEVIVHP 248
QY 242 OSIYHAWGFCDGGLMAHLGPAIDRHAIGFALNMPGREGVAVARIDLAQISLTFORPDE 301
Db 249 OSIYHAWGYTDGSLVALQOLGCPDMRTAIGALSTPKRCDLVERLDFAIRLDEEAPDE 308
QY 302 ERPALRLARDVMAARGLSGAFAFNAKEIALDHFIAGRIGFLDMAAAVEETLAGVSTDP 361
Db 309 VRPAIKARARAMEGGVQGAVALNGAKRTALDAFIKGRIGFLMAAEIVEXYMDGLAGLP 367
QY 362 FGKVPDALEEVLANDHARAAAE 365
Db 368 ---AATSMVDVFAADERARRAAAE 388

```

RESULT 2

```

ID DNR_AGR5 STANDARD; PRT; 397 AA.
AC Q8UC86;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 1-DEOXY-D-XYLULOSE 5-PHOSPHATE REDUCTOISOMERASE (EC 1.1.1.267) (DXP
reductoisomerase) (1-deoxyxylulose-5-phosphate reductoisomerase).
GN DNR OR ATU2612 OR AGR_C_4736.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; Pubmed=11743193;
RA Wood D.W., Setudal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutyavin T., Levy R., Li M.-Y., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.,
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RT Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; Pubmed=11743194; Miller N., Blanchard M.,
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmlet K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Woumler C., Allinger M., Doughly D., Scott C., Iappas C., Markelz B.,
RA Planagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.,
RT Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RT Science 294:2323-2328(2001).
CC -1- FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction
CC of 1-deoxy-D-xylulose-5-phosphate (DXP) to 2-C-methyl-D-erythritol
CC 4-phosphate (MEP) (By similarity).

```

```

CC -1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADP(+)
CC = 1-deoxy-D-xylulose 5-phosphate + NADPH.
CC -1- PATHWAY: Nomevalonate terpenoid biosynthesis pathway; second
CC step.
CC -1- SIMILARITY: BELONGS TO THE DNR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE009208; AAL43593.1; -.
DR EMBL: AE008174; AAR68334.1; -.
KW Isoprene biosynthesis; Oxidoreductase; NADP; Complete proteome.
FT NP_BIND 14 21 NADPH (POTENTIAL).
SQ SEQUENCE 397 AA; 42303 MW; 0A24C408C2D47094 CRC64;

```

Query Match 50.1%; Score 994; DB 1; Length 397;

Best Local Similarity 50.4%; Pred. No. 5.5e-63; Matches 197; Conservative 64; Mismatches 126; Indels 4; Gaps 1;

```

QY 2 RSLIFGATGSGTSTLDVIERLGRDFEIALTGNNINILAEQARRIGAEIATVAD 61
Db 9 RRLTILSTGSGTSTLDVIERLGRDFEIALTGNNINILAEQARRIGAEIATVAD 68
QY 62 DCLPALREALAGTGEVAGGAOATAEADRPADMTMSAIVAGACIVPMRLKRGRTAL 121
Db 69 DRYELKDAISGSGLEVAAGRSGLTEAERDAGWMAIVAGGFLAARRKADIAL 128
QY 122 ANKESLVATAGOLMRTAENGATILPVDESHSAFVALAGEDTACVERVITTAAGGPPRD 181
Db 129 ANKCVISAGSLFIDVAEAGGRLIPVDESHALFOVLENOGRHVAEIVITLASSGPPRT 188
QY 182 WSLERIRACTYAEAOAHPMNSMGORISIDSAMFNKALELIEFEFGFEPDRIEAVVHP 241
Db 189 WSRPMSVNTADIRAHPMNSMGKISIDSAMFNKALEMIEARHLFRLEPQIEVIVHP 248
QY 242 OSIYHAWGFCDGGLMAHLGPAIDRHAIGFALNMPGREGVAVARIDLAQISLTFORPDE 301
Db 249 OSIYHAWGYTDGSLVALQOLGCPDMRTAIGALSTPKRCDLVERLDFAIRLDEEAPDE 308
QY 302 ERPALRLARDVMAARGLSGAFAFNAKEIALDHFIAGRIGFLDMAAAVEETLAGVSTDP 361
Db 309 ARPALRLARALREGGVQGAVALNGAKRTALDAFIKGRIGFLMAAEIVEXYMDGLAGLP 367
QY 362 FGKVPDALEEVLANDHARAAAEARAAAE 392
Db 366 -GRTAETMDVFAADERARRAAAE 395

```

RESULT 3

```

ID DNR_CAUCR STANDARD; PRT; 399 AA.
AC Q9A709;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 1-DEOXY-D-XYLULOSE 5-PHOSPHATE REDUCTOISOMERASE (EC 1.1.1.267) (DXP
reductoisomerase) (1-deoxyxylulose-5-phosphate reductoisomerase).
GN DNR OR CCI1917.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; Pubmed=11259647;
RA Niernan W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,

```


RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA Deboy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of *Caulobacter crescentus*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 CC -1- FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction
 of 1-deoxy-D-xylulose-5-phosphate (DXP) to 2-C-methyl-D-erythritol
 4-phosphate (MEP) (by similarity).
 CC -1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADP(+) ->
 CC -1- 1-deoxy-D-xylulose 5-phosphate + NADPH.
 CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second
 CC step.
 CC -1- SIMILARITY: BELONGS TO THE DXR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AE005865; AAK23892.1; -
 DR TIGR: CC1917; -
 DR InterPro: IPR003821; DXP_reductoisomerase.
 DR Pfam: PF02670; DXP_reductoisom. 1.
 DR TIGRFAMS: TIGR00243; Dxr: 1.
 KW Isoprene biosynthesis; Oxidoreductase; NADP; Complete proteome.
 FT NP_BIND 13 20 NADPH (POTENTIAL).
 SQ SEQUENCE 399 AA; 41456 MW; B68D94CD0A47C25 CRC64;

Query Match 45.3%; Score 899.5; DB 1; Length 399;
 Best Local Similarity 48.6%; Pred. No. 2.5e-56;
 Matches 191; Conservative 52; Mismatches 147; Indels 3; Gaps 2;

OY 2 RSLSTFGATSGTGEFTDVMRKSGPEAFRTVALTNGRNIRRLAEMARALKELAYTAHE 61
 DB 8 RRVVVGSTGSLSTLSLFEEGAP--VQILALTGCRNVERIEGARRRMKPSLAVIDE 65
 OY 62 DCIPALREALAGTGTAVAGAAQIAEAADRPADMTSAIVGAAGLVGMRALKHGTAL 121
 DB 66 SRIDDLRAELAGTGTAVAGAAQIAEAADRPADMTSAIVGAAGLVGMRALKHGTAL 125
 OY 122 ANKESLVTAAGQILMTAENGATILPVDESHSAVFOALAGEPTACERYIITASGGPFD 181
 DB 126 ANKESLVTAAGQILMTAENGATILPVDESHSAVFOALAGEPTACERYIITASGGPFD 185
 OY 182 WSLERIRACTVAEAQAHPMWMSGORISIDSASFNFKALELIERREFGEPDRIEAVHP 241
 DB 186 WDKAAMARATPBGALIHHPWMSMGAKISVDSATPMNKGLEIETSYFAFPEDPDVDVHP 245
 OY 242 OSIIVHAMVFCDDGLMAHLGAPADMRAHIGFALNMPGRGEPVAVARIDLAQIASLTFOKPDE 301
 DB 246 OSIIVHAMVFCDDGLMAHLGAPADMRAHIGFALNMPGRGEPVAVARIDLAQIASLTFOKPDE 305
 OY 302 EEPFALRLADVMAARGISGAAPNAKETAALDHFIAGRIIGFLDMAVVEETAGV-STDP 360
 DB 306 EEPFALRLADVMAARGISGAAPNAKETAALDHFIAGRIIGFLDMAVVEETAGV-STDP 365
 OY 361 LFGKVPDALEEVYALMDHLARRAAEEAAGLRQK 393
 DB 366 LSYAESDAVETAMLDIGSARRTAAYVAOKROR 398

RESULT 4
 ID DXR_ZYMO STANDARD: PRT; 388 AA.
 AC O9XSF2: O9RIA9:
 DT 30-MAY-2000 (rel. 39, Created)
 DT 16-OCT-2001 (rel. 40, Last sequence update)
 DT 15-JUN-2002 (rel. 41, Last annotation update)

DE 1-deoxy-D-xylulose 5-phosphate reductoisomerase (EC 1.1.1.267) (DXP
 DE reductoisomerase) (1-deoxyxylulose-5-phosphate reductoisomerase).
 GN DXR.
 OS Zymomonas mobilis.
 OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;
 OC Zymomonas.
 OX NCBI_TaxID=542;
 RN [1]
 RP SEQUENCE FROM N.A. AND CHARACTERIZATION.
 RC STRAIN=ATCC 31821 / ZM4 / CP4;
 RX MEDLINE=20461176; PubMed=11004410.
 RA Grollie S., Bringer-Meyer S., Salm H.;
 RT "Isolation of the dxr gene of *Zymomonas mobilis* and characterization
 RT of the 1-deoxy-D-xylulose 5-phosphate reductoisomerase";
 RL FEMS Microbiol. Lett. 191:131-137(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 31821 / ZM4 / CP4;
 RA Lee H.J., Kang H.S.;
 RT Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction
 of 1-deoxy-D-xylulose-5-phosphate (DXP) to 2-C-methyl-D-erythritol
 4-phosphate (MEP).
 CC -1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADP(+) ->
 CC -1- 1-deoxy-D-xylulose 5-phosphate + NADPH.
 CC -1- COFACTOR: REQUIRES A DIVALENT CATION; MAGNESIUM, MANGANESE OR
 CC COBALT.
 CC -1- ENZYME REGULATION: INHIBITED BY POSITRONIUM.
 CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second
 CC step.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SIMILARITY: BELONGS TO THE DXR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AJ250714; CAB60758.1; -
 DR EMBL: AF124757; AAD29659.1; -
 DR InterPro: IPR003821; DXP_reductoisomerase.
 DR Pfam: PF02670; DXP_reductoisom. 1.
 DR TIGRFAMS: TIGR00243; Dxr: 1.
 KW Isoprene biosynthesis; Oxidoreductase; NADP; Magnesium; Manganese;
 KW Cobalt.
 FT NP_BIND 10 17 NADPH (POTENTIAL).
 FT CONFLICT 226 226 F -> Y (IN REF. 2).
 SQ SEQUENCE 388 AA; 41842 MW; 2E161B712089613F CRC64;

Query Match 44.1%; Score 875; DB 1; Length 388;
 Best Local Similarity 46.9%; Pred. No. 1.3e-54;
 Matches 183; Conservative 67; Mismatches 130; Indels 10; Gaps 3;

OY 2 RSLSTFGATSGTGEFTDVMRKSGPEAFRTVALTNGRNIRRLAEMARALKELAYTAHE 61
 DB 5 RYTVLGATSGTGHSTLTLIER--NDRYQVALLTANRWKDKLADAKRTNARAVIADP 62
 OY 62 DCIPALREALAGTGTAVAGAAQIAEAADRPADMTSAIVGAAGLVGMRALKHGTAL 121
 DB 63 SLYNDKELAGSVEAAGADALVAAAMGADMTAAIIGCAGLKATLAIRKGTVAL 122
 OY 122 ANKESLVTAAGQILMTAENGATILPVDESHSAVFOALAGEPTACERYIITASGGPFD 181
 DB 123 ANKESLVTAAGQILMTAENGATILPVDESHSAVFOALAGEPTACERYIITASGGPFD 182
 OY 182 WSLERIRACTVAEAQAHPMWMSGORISIDSASFNFKALELIERREFGEPDRIEAVHP 241
 DB 183 TSLAEMATVTPERAVOHPMWMSMGAKISVDSATPMNKGLEIETSYFAFPEDPDVDVHP 242
 OY 242 OSIIVHAMVFCDDGLMAHLGAPADMRAHIGFALNMPGRGEPVAVARIDLAQIASLTFOKPDE 301

```

Db 243 QSVHSMWEYLDGSILOIGSPDMKRTPIGHTLAMPKREMTAESIDETKLOMDFEADY 302
Oy 302 ERPALRLARDVMAARGLSCAFAFNAREIALDHFIAIGIGFLDMAAVEETLSTPL 361
Db 303 ERFPALRLAMSTISGSGARRPVMNAANEIYAVALDKKIGFLDIKIVYKTL-----DHY 357
Oy 362 FGKVPDALEEVYLANDHLARRAAEAGLRQ 391
Db 358 TPATPSSLEDFVAIDNEAR---IOAAALME 384

RESULT 5
DXR_RALSO STANDARD; PRT; 393 AA.
ID DXR_RALSO STANDARD; PRT; 393 AA.
AC 08X215;
DT 15-JUN-2002 (Rel. 41, last sequence update)
DT 15-JUN-2002 (Rel. 41, last sequence update)
DE 1-deoxy-D-xyulose 5-phosphate reductoisomerase (EC 1.1.1.267) (DXP
reductoisomerase) (1-deoxyxyulose-5-phosphate reductoisomerase).
GN DXR OR RSC1410 OR RS05282.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group.
OX NCBI_Taxid=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangelot S.,
RA Ariat M., Billault A., Brotter P., Camus J.C., Cattolico L.,
RA Chandler C., Lavie N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin G., Moisan A., Robert C., Saurin W., Schlex T.,
RA Siguler P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
CC -1- FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction
of 1-deoxy-D-xyulose-5-phosphate (DXP) to 2-C-methyl-D-erythritol
4-phosphate (MEP) (By similarity).
CC -1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate +
-1-deoxy-D-xyulose 5-phosphate + NADPH.
CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second
step.
CC -1- SIMILARITY: BELONGS TO THE DXR FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC EMBL: AL646064; CAD15112.1;
DR InterPro: IPR003821; DXP_reductisomase.
DR Pfam: PF02670; DXP_reductoisom; 1.
DR TIGRPFAMs: TIGR00243; Dxr; 1.
KW Isoprene biosynthesis; Oxidoreductase; NADP; Complete proteome.
FT NF_BIND 7 NADPH (POTENTIAL).
SQ SEQUENCE 393 AA; 41763 MW; 5A995A7C8FF32BED CRC64;

Query Match 40.2%; Score 797.5; DB 1; Length 393;
Best Local Similarity 43.6%; Pred. No. 3.6e-49;
Matches 170; Conservative 73; Mismatches 134; Indels 13; Gaps 5;
Oy 1 MRSISITGAGSGESTFEDLVMRKGGPEARFRTVALTGGRNIRRLAEARALKALAVTAH 60
Db 1 MMRITVIGAGSGISTDPLDVRRH--PDYRVFALFANTQVDKIALCVRFRPAAVAVGS 58
Oy 61 EDCIAPALRELA--GCTGEVAGAGQALAEADNR-ADMTSATYAGAGLVPGMALKHGR 117

```

```

Db 59 ATAAEVLNDQGAERTGIDIRFGEPALEEAAPHDCAVMAIYVAGAGLRPTLAVRAGK 118
Oy 118 TLALANKESVITAGQLMRTAQENGATILPVDSHSAVFQALAGEDTAC---VERVITTA 174
Db 119 RVLANKREALVMSGLAFEDAYRQHGATYLPIDSEHNAIFQCLPQORPSFGVGARIYVTA 178
Oy 175 SGGEPRDMSLEIRACTYVAEQAHPNMSGQRISIDSAFMENKALELLETRFEGFEPDR 234
Db 179 SGCFPRTRAVETLAEVTPDQACAPHNWVGRIKISVDSATMANKGLEVEAHMFLNVEVER 238
Oy 235 IEAVVHPQSYVHAWVFCDDGLMAHLGPRADMRHAIIGFALNMPGGEVPAVIAIQAASL 294
Db 239 LEVLIHPQSYVHSMVAVDGSLVLAQLGNPDRTPIAGLAPXETIEGVPLDLAATGTL 298
Oy 295 TFQKDEERFPALRLARDVMAARGLSCAFAFNAREIALDHFIAIGIGFLDMAAVEETLA 354
Db 299 AFEAPDLHRFPCLALFALAGTAPRVANAEVNEARLQGRIRRTETIAAVGDTLA 358
Oy 355 GVSTDPLEFGKVPDALEEVYLANDHLARRAAE 384
Db 359 RTAIGP-----ADSIDTVEFAADQAARRRAE 383

RESULT 6
DXR_PSEAE STANDARD; PRT; 396 AA.
ID DXR_PSEAE STANDARD; PRT; 396 AA.
AC 09KGD6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, last sequence update)
DT 15-JUN-2002 (Rel. 41, last annotation update)
DE 1-deoxy-D-xyulose 5-phosphate reductoisomerase (EC 1.1.1.267) (DXP
reductoisomerase) (1-deoxyxyulose-5-phosphate reductoisomerase).
GN DXR OR PA3650.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_Taxid=287;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=20487113; PubMed=11034300;
RA Altincelik B., Hintz M., Sanderbrand S., Wiesner J., Beck E.,
RA Jomaia H.;
RT "Tools for discovery of inhibitors of the 1-deoxy-D-xyulose
5-phosphate (DXP) synthase and DXP reductoisomerase: an approach with
enzymes from the pathogenic bacterium Pseudomonas aeruginosa.";
RL FEMS Microbiol. Lett. 190:329-333(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15692 / PAOI;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAOI, an
opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -1- FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction
of 1-deoxy-D-xyulose-5-phosphate (DXP) to 2-C-methyl-D-erythritol
4-phosphate (MEP).
CC -1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADP(+)
-1-deoxy-D-xyulose 5-phosphate + NADPH.
CC -1- ENZYME REGULATION: INHIBITED BY POSITROMYCIN AND 3-(N-ACETYL-N-
HYDROXYAMINO)-PROPYLPHOSPHONIC ACID (PR-900098).
CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second
step.
CC -1- SIMILARITY: BELONGS TO THE DXR FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: AF282879; AAF97241.1; -
DR EMBL: AE004785; AAG07038.1; -
DR InterPro: IPR003821; DXR_rediisomerase.
DR Pfam: PF02670; DXR_reductoisom; 1.
DR TIGRFAMS: TIGR00243; DXR: 1.
KW Isoprene biosynthesis; Oxidoreductase; NADP; Complete proteome.
FT NP_BIND 10 NADPH (POTENTIAL).
SQ SEQUENCE 396 AA: 42508 MW; 2ECB9D8A94A9CAD7 CRC64;

Query Match 39.1%; Score 775.5; DB 1; Length 396;
Best Local Similarity 42.8%; Pred. No. 1.3e-47;
Matches 166; Conservative 73; Mismatches 134; Indels 15; Gaps 5;

OY 4 LSIPTGATSGISGTEPDLVNRKGGPEAFRTVALTGGNIRRLAEMARLKAELAVTAHEDC 63
DB 7 ISVLATGSGISGLTDVQGRH--PDREAFALTFGRSLAELALCLRHPRVAVPEQAA 64
OY 64 LPALREALAGT--TEVAGGAQAIAEADRP-ADMTSAIYVAGAGIVPGMRALKHGRTIA 120
DB 65 AATAGSILAAAGIRTRVLEGEQALCEVASAPEDVMAAIVAGAGIPSTLAVAGKRYL 124
OY 121 LANKESLVYAGOLMRTQENGATLIPVDSHSAVFQAL-----GEONACVERVITTS 175
DB 125 LANKELVWAGLAFMOAVRSGAVLLPIDSEHNAIFQSLPRRYADCLEVGVGRILLTTS 184
OY 176 GGFPRDMSLERIRACTVAAQAHPNMWMOGRISIDSASFNRKALELIERREFGEPDRI 235
DB 185 GGPFRETPLEQLASTYPEQACAHPMWMOGRKISVDSASWNNKGLIEACWFLDQPSGV 244
OY 236 EAVVHPQSVIHAMVGFCDGLMAHLGPRADMRHAIQFALNMPGRGEVPAVIDLAQIASLT 295
DB 245 EVVHHQSVIHAMVGVDSVTAQGNPDMRPISTAYMAWMPRIDSGVSLPMFAVGRAD 304
OY 296 FOKPDEERPALRLARDVMAARGLSGAANAKETALDHFIAIGRIGFLDMAVVEETLAG 355
DB 305 FOKPDEORPCRLRLASQAETGSGAPAMLNANEAVAALFRRHFRISDIAYIIEVLR 364
OY 356 VSTDPLFGKVPDALEVLAMDLARRAA 383
DB 365 EAVTAV-----ESLDQVLAADRRARSVA 387

RESULT 7
DXR_NEIMB
ID DXR_NEIMB STANDARD: PRT; 394 AA.
AC 09K1G8;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE 1-deoxy-D-xylose 5-phosphate reductoisomerase (EC 1.1.1.267) (DXP
reductoisomerase) (1-deoxyxylulose-5-phosphate reductoisomerase).
GN DXR OR NMB0184.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN (1)
RP SEQUENCE FROM N. A.
RC STRAIN-MC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., Deboy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
RA Cotton M.D., Ullrichback T.R., Khouli H., Qin H., Yamatevan J.,
RA Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Koxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain

RT MC58.";
RL Science 287:1809-1815(2000).
CC -! FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction
CC of 1-deoxy-D-xylose-5-phosphate (DXP) to 2-C-methyl-D-erythritol.
CC 4-phosphate (MEP) (by similarity).
CC -! CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADP(+)
CC = 1-deoxy-D-xylose 5-phosphate + NADPH.
CC -! PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second
CC step.
CC -! SIMILARITY: BELONGS TO THE DXR FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: AE002375; AAF40641.1; -
DR TIGR: NMB0184; -
DR InterPro: IPR003821; DXR_rediisomerase.
DR Pfam: PF02670; DXR_reductoisom; 1.
DR TIGRFAMS: TIGR00243; DXR: 1.
KW Isoprene biosynthesis; Oxidoreductase; NADP; Complete proteome.
FT NP_BIND 9 16 NADPH (POTENTIAL).
SQ SEQUENCE 394 AA: 41920 MW; AA85397E5EB7BD2E CRC64;

Query Match 37.8%; Score 750.5; DB 1; Length 394;
Best Local Similarity 44.0%; Pred. No. 7.3e-46;
Matches 173; Conservative 55; Mismatches 150; Indels 15; Gaps 6;

OY 4 LSIPTGATSGISGTEPDLVNRKGGPEAFRTVALTGGNIRRLAEMARLKAELAVTAHEDC 63
DB 6 LTIISTGSGISTLDVSRH--PEKFRVFLAGHGVKLAACCTHETAVVAADAH 63
OY 64 LPALREALA--GTGEVAGGAQAIAEADRP-ADMTSAIYVAGAGIVPGMRALKHGRTIA 120
DB 64 AARLEALLKRDGTATQVHGAQALVDVADESGVWCIVAGVGLPSLAAAKGRTIY 123
OY 121 LANKESLVYAGOLMRTQENGATLIPVDSHSAVFQAL-----AGE-DTACVERVITTS 175
DB 124 LANKETLVVSGALFMETARANGAAVLLPIDSEHNAVFQVLPDYAGLNLHGASITLTS 183
OY 176 GGFPRDMSLERIRACTVAAQAHPNMWMOGRISIDSASFNRKALELIERREFGEPDRI 235
DB 184 GGPFRLADINTDRITPAQAVKHPNMWMOGRKISVDSATMKNKGLIEAHMLPNCPPDL 243
OY 236 EAVVHPQSVIHAMVGFCDGLMAHLGPRADMRHAIQFALNMPGRGEVPAVIDLAQIASLT 295
DB 244 EVVHHQSVIHAMVGRDQSVLAQGNPDMRPISTAYMAWMPRIDSGVSLPMFAVGRAD 303
OY 296 FOKPDEERPALRLARDVMAARGLSGAANAKETALDHFIAIGRIGFLDMAVVEETLAG 355
DB 304 FOKPDEORPCRLRLAYEANNAGAPCVLNANEAVAALFDGQIKFTDIARTVAHCLAQ 363
OY 356 VSTDPLFGKVPDALEVLAMDLARRAAEAG 388
DB 364 DFDGTL-----GDIGGLAQDARTRAQAARAFIG 391

RESULT 8
DXR_NEIMA
ID DXR_NEIMA STANDARD: PRT; 394 AA.
AC 09UX33;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 41; Last sequence update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE 1-deoxy-D-xylose 5-phosphate reductoisomerase (EC 1.1.1.267) (DXP
reductoisomerase) (1-deoxyxylulose-5-phosphate reductoisomerase).
GN DXR OR NMA0083.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.

OX NCBI_TaxID-65699;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX STRAIN-22491 / Serogroup A / Serotype 4A;
 RA MEDLINE-20222556; PubMed-10761919;
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
 RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
 RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holtroyd S.,
 RA Jagers K., Leather S., Moule S., Mungall K., Quail M.A.,
 RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
 RA Whitehead S., Spratt B.G., Barrall B.G.;
 RA "complete DNA sequence of a serogroup A strain of *Neisseria*
 RT meningitidis 22491.";
 RL Medline 404:502-506(2000).
 CC -1- FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction
 CC of 1-deoxy-D-xylulose-5-phosphate (DXP) to 2-C-methyl-D-erythritol
 CC 4-phosphate (MEP) (By similarity).
 CC -1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADP(+)
 CC -1- -1-deoxy-D-xylulose 5-phosphate + NADPH.
 CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second
 CC step.
 CC -1- SIMILARITY: BELONGS TO THE DXR FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AL162752; CAB83399.1;
 DR InterPro: IPR003821; DXP_reductoisomerase.
 DR Pfam: PF02670; DXP_reductoisom; 1.
 DR TIGRFAMs: TIGR00243; Dxr; 1.
 KM Isoprene biosynthesis; Oxidoreductase; NADP: Complete proteome.
 FT NP_BIND 9
 SO SEQUENCE 394 AA; 41912 MW; B5D361B224806D8 CRC64;
 Query Match 37.5%; Score 743.5; DB 1; Length 394;
 Best Local Similarity 43.0%; Pred. No. 2.3e-45;
 Matches 169; Conservative 58; Mismatches 151; Indels 15; Gaps 5;
 OY 4 LSTIGANGSGESTFDLMKRGKGEAFRTVALTGCGRNTRKAEAKALKAFLATTAHEDC 63
 DB 6 LTIIGSTGSGESTFLDVVSRH--PEKFRVFALAGHKQVEKLAACQGFHPEYAVVADAEN 63
 OY 64 LPALREALA--GTGTEVAGGAQAIAEAADR-PADMTSAIVGAGLVPGMALKHGRTIA 120
 DB 64 AARLEALLKRGATVQLHGAQALVDVAASADEVSGVCAITGAVGLSALAAOKKTIY 123
 OY 121 LANKESLVTAAGQLMKRAQENGATILPVSESAVFOALAGEDTA-----CVERVIITAS 175
 DB 124 LANKETLVVSGALFEMETARANGAVALPVDSSEHNAVFOVLPRDYGRINENHGIASITLITAS 183
 OY 176 GGPRDMSLEIRACTYAAEAQAHNMSCORISIDSAMFKALELLETREFFSEPDRI 235
 DB 184 GGPELVADLNTFDSITFDQAVKHPNMRGRISVDSATMMKRGLELLEAHLFNCPPDKL 243
 OY 236 EAVVHPISIVAVAGFGDGLMAHLGPADMRHAGFLNMPGGEVAVPARIDLAQIASLT 295
 DB 244 EYVHPISIVHSVRYRQDGSVLAOLGNPDMRTPLATYCLGPERIDSGVGLDIFALSALT 303
 OY 296 FOKPDEERFALRLARVMAAGLSGAFAVNAKEIALDHFAGRIIGFLDMAVAVEETLAG 355
 DB 304 FOKPDEFORFOLKAYEMANAGAAPCVLNANAEAAVAALFDGQIKFTDIAKTVAHCLSQ 363
 OY 356 VSTDPLFGKVPDALEEVLANDHILARRAAEEELAG 388
 DB 364 DFDSDGI-----GDIGGLADADARTRAARAFITG 391
 RESULT 9

DXR_ARATH
 ID DXR_ARATH STANDARD; PRT; 477 AA.
 AC O9XFS9: O9M6U2;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 1-deoxy-D-xylulose 5-phosphate reductoisomerase, chloroplast precursor
 DE (RC 1.1.1.267) (DXP reductoisomerase) (1-deoxyxylulose-5-phosphate
 DE reductoisomerase).
 GN DXR OR AT5G62790 OR MGB2.11.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RC SEQUENCE FROM N.A.
 RA STRAIN-cv. Columbia;
 RA Campos N., Lois L.M., Cunillera N., Carretero L., Alameda I.,
 RA Hoefli J.-F., Pale-Grosdemange C., Rohmer M., Ferrer A., Boronat A.;
 RT "Isolation and characterization of a cDNA from *Arabidopsis thaliana*
 RT encoding 1-deoxy-D-xylulose 5-phosphate reductoisomerase, the first
 RT committed enzyme of the non-mevalonate pathway for isoprenoid
 RT biosynthesis.";
 RT Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RC SEQUENCE FROM N.A.
 RA STRAIN-cv. Columbia;
 RX MEDLINE-98290546; PubMed-9628582;
 RA Sato S., Kaneko T., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,
 RA Tabata S.;
 RA "Structural analysis of *Arabidopsis thaliana* chromosome 5. IV.
 RT Sequence features of the regions of 1,456,315 bp covered by nineteen
 RT physically assigned P1 and TAC clones.";
 RL DNA Res. 5:41-54(1998).
 RN [3]
 RP SEQUENCE OF 72-477 FROM N.A.
 RC STRAIN-cv. Columbia;
 RA Schwender J., Mueller C., Zeidler J., Lichtenthaler H.K.;
 RT "Cloning and heterologous expression of a cDNA coding for 1-deoxy-D-
 RT xylulose-5-phosphate reductoisomerase of *Arabidopsis thaliana*.";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction
 CC of 1-deoxy-D-xylulose-5-phosphate (DXP) to 2-C-methyl-D-erythritol
 CC 4-phosphate (MEP) (By similarity).
 CC -1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADP(+)
 CC -1- -1-deoxy-D-xylulose 5-phosphate + NADPH.
 CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second
 CC step.
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- SIMILARITY: BELONGS TO THE DXR FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AF148852; AAF73140.1;
 DR EMBL: AB009053; BAB10848.1;
 DR EMBL: AJ242588; CAB43344.1;
 DR InterPro: IPR003821; DXP_reductoisomerase.
 DR Pfam: PF02670; DXP_reductoisom; 1.
 DR TIGRFAMs: TIGR00243; Dxr; 1.
 KM Isoprene biosynthesis; Oxidoreductase; NADP: Chloroplast;
 KW Transic peptide.
 FT TRANSIT 1
 FT CHAIN 1
 FT NP_BIND 86 93
 FT FT
 SO SEQUENCE 477 AA; 51963 MW; C06A455AB73ACA7C CRC64;
 CHLOROPLAST (POTENTIAL).
 1-DEOXY-D-XYLULOSE 5-PHOSPHATE
 REDUCTOISOMERASE.
 NADPH (POTENTIAL).

[illegible]

DE reductoisomerase) (1-deoxyxylulose-5-phosphate reductoisomerase).
 GN DNR OR 20184 OR ECS0175.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_Taxid=83334;
 RN
 RP
 RC STRAIN=O157:H7 / EDI933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lam A., Dimantanta E.T., Potamoukis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen J., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 RL Nature 409:529-533(2001).
 RN [2]
 RP
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=1128796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsuno E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kihara S., Shida T., Hattori M., Shunagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12";
 RL DNA Res. 8:11-22(2001).
 CC
 CC -1- FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction
 of 1-deoxy-D-xylulose-5-phosphate (DXP) to 2-C-methyl-D-erythritol
 4-phosphate (MEP) (By similarity).
 CC
 CC -1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADP(+)
 -> 1-deoxy-D-xylulose 5-phosphate + NADPH.
 CC
 CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second
 step
 CC
 CC -1- SIMILARITY: BELONGS TO THE DXR FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AE005193; AAC54475.1; -;
 CC EMBL: AP002550; BAB3598.1; -;
 CC DR InterPro: IPR003821; DXP_reductoisomase.
 CC DR Pfam: PF02670; DXP_reductoisom; 1.
 CC KW Isoprene biosynthesis; Oxidoreductase; NADP; Complete proteome.
 CC FT NS_BIND 7
 CC FT SEQUENCE 398 AA; 43361 MW; 8B532683A4FEF082E CRC64;
 SQ
 Query Match 36.9%; Score 733; DB 1; Length 398;
 Best Local Similarity 40.7%; Pred. No. 1.3e-44;
 Matches 162; Conservative 72; Mismatches 144; Indels 20; Gaps 5;
 QY 1 MRSLSITGATGSGTGFEDLVMRKRGPEAFRTVALTGTGRNIRLAEMARALKAELATYAH 60
 DB 1 MKQITIGSGTICGSLDYV--RHNEHPRVVALVAGKNTRAVEQCLESPRYAVMD 58
 QY 61 EDCPLALREAL--AGTTEVAGGAQAIAB--ADRPADWTSAIVGAAGLVGKRALHGR 117
 DB 59 EASAKLKTMLQOGSRTVELSGQAACDMALEDDVDVMAIYGAAGLLPTLAIRAGK 118
 QY 118 TLALANESLVTAAGQLMRRAQENGATILPYDSEHSAVFOAL-----AGEDTACV 167
 DB 119 TILANESLVTCGRLEMDVAVKOSKQALLPDSSEHNAIFOSLPPIOHNLGADLEONGV 178
 QY 168 ERVITTSAGGPRDMSERIRACTVAEAOAHPNMSMGORISIDASFNKALEIEREF 227
 DB 179 VSIILTSGGPRFETPLRLDLATMTPDQACRHPNMSMGKRISVDSATMANKLEYIEARWL 238

QY 228 GFEPDRIEAVVHPOSIVHANWGCDCGLMAHLCPADMRHAIIGFALMWPGRGEVPAVID 287
 DB 239 FNASQMEVLIHPQSVTHSWRVQDGSVYLAQLGEPDMRFPFIATMTAMPNVNNGVAPLD 298
 QY 288 LAQIASLTFQKPPDEERPALRLADVMAARGLSCAATNAKEILDLFIAGRIEFLDMAA 347
 DB 299 FCKLSALTFEAPDYDRYPCIKLMEAEQGAATTAANAETVVAFLAQOIRFTDIAA 358
 QY 348 VVEETLAGVSDPLFGVDPALVEELAMDHLARRAEE 385
 DB 359 LNLSTL-----EKMDREPCQVDVLSVDSAREVARK 391
 RESULT 12
 DNR_ECOLI
 ID DNR_ECOLI STANDARD; PRT; 398 AA.
 AC P45568; P77209;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 1-deoxy-D-xylulose 5-phosphate reductoisomerase (EC 1.1.1.267) (DXP
 DE reductoisomerase) (1-deoxyxylulose-5-phosphate reductoisomerase).
 GN DNR OR B0173.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_Taxid=562;
 RN [1]
 RP
 RC STRAIN=K12 / W3110;
 RX MEDLINE=98374274; PubMed=9707569;
 RA Takahashi S., Kuzuyama T., Matanabe H., Seto H.;
 RT "A 1-deoxy-D-xylulose 5-phosphate reductoisomerase catalyzing the
 RT formation of 2-C-methyl-D-erythritol 4-phosphate in an alternative
 RT nonmevalonate pathway for terpenoid biosynthesis";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:9879-9884(1998).
 RN [2]
 RP
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12";
 RL Science 277:12453-12474(1997).
 RN [3]
 RP
 RC STRAIN=K12 / W3110;
 RX Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
 RA Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,
 RA Mizuno T., Makino K., Nakata A., Yura T., Sempel G., Mizobuchi K.;
 RT "Systematic sequencing of the Escherichia coli genome: analysis of the
 RT 4.0-6.0 min (189,987 - 281,416bp) region";
 RL Submitted (Feb-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP
 RC STRAIN=K12 / W3110;
 RX Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
 RA Davis K., Federspiel N., Hyman R., Kalman S., Komp C., Kurli O.,
 RA Laskhari D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,
 RA Davis R.W.;
 RT Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP
 RC STRAIN=K12 / W3110;
 RX MEDLINE=93077430; PubMed=1447125;
 RA Yamana K., Ogura T., Miki H., Hiraga S.;
 RT "Identification and characterization of the smba gene, a suppressor
 RT of the mukB null mutant of Escherichia coli";
 RL J. Bacteriol. 174:7517-7526(1992).
 RN [6]
 RP
 RC STRAIN=K12 / W3110;
 RX IDENTIFICATION
 MEDLINE=96032851; PubMed=7567469;

[illegible]

CC 4-phosphate (MEP) (By similarity).
CC -1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADP(+)
CC = 1-deoxy-D-xylulose 5-phosphate + NADPH.
CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second
CC step.
CC -1- SIMILARITY: BELONGS TO THE DXR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE006235; AAK04072.1; -
DR InterPro: IPR003821; DXR_red isomerase.
DR Pfam: PF02670; DXR_red isom; 1.
DR TIGRfams: TIGR00243; Dxr; 1.
KW Isoprene biosynthesis; Oxidoreductase; NADP; Complete proteome.
FT NP_BIND 13 20 NADPH (POTENTIAL).
SQ SEQUENCE 405 AA; 44207 MW; 1CFF7B44847F8BF7 CRC64;

Query Match 36.1%; Score 716; DB 1; Length 405;
Best Local Similarity 39.3%; Pred. No. 2e-43;
Matches 156; Conservative 74; Mismatches 147; Indels 20; Gaps 5;

QY 1 MRSLSIFGATGSGTSGTSTPVMKSGPEAFRTVALTGCGRNIRRLAEMARALKAEALVTAH 60
DB 7 MKRIYILGSGTSGTSTLSVTH--NPKYQVAFVAGRNVELMFQOCLTFQPSFALDD 64

QY 61 EDCLPALREALAG--TGTEVAGCAQAIABADRP--ADWTMSAIVGAGLVPGMRALKHGR 117
DB 65 DVAAMKLAELKAKHOSQTVLAGQAICELAHPEADMMAIVGAGLPTLSAYKAKR 124

QY 118 TLALANKESLVTAQQLMRTAQENGATILPVDESHSAVFQALAGE-----DTACV 167
DB 125 RVLANKKALVTCGOLFIDAVRESQQLPVDSEHNAIFOSLPPEAQRQIGCPILSELGI 184

QY 168 ERVITTAGSGPFRDMSLERIRACTVAEAQAHPNWSMGORISIDSAMFNKALETITREF 227
DB 185 SKIVLGSQGPFRYPPEQFOITPQAQAVAHHPWSMGKRISVDSATMNNKGLEYIEARWL 244

QY 228 FGFEEDRIEAVVPOISIVAMVFCDDGLMAHIGPADMRHAIGFALNMPGRGCEVPVARI 287
DB 245 FNASAEEMEVIIHPQSIHSMRYIDGSVIAQMGNDMRTPIAETMAYPSRTVAGYEPID 304

QY 288 LAQIASLTFQKPDDEERFALRLARDVMAARGLSGAENAKETALDHTAGRIGFLDMAA 347
DB 305 FYQNLGTFIEPDYQYPCIKLAIDAFSAGQYATTMNAANEIAVASFLDNKIKFTDIAR 364

QY 348 VVEETLAGVSTDPFGKVPDALAEVYLAAMDHLARRAAE 384
DB 365 LMQLVYSKIQPOKI-----HCIEDVLEVDKKARELSQ 396

Search completed: April 15, 2003, 13:57:00
Job time : 26 secs

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2003, 13:56:29 ; Search time 88 Seconds
(without alignments)
922.529 Million cell updates/sec

Title: US-09-673-198-30

Perfect score: 1984

Sequence: 1 MRSLSIFGATSGISGSPFDL.....MDHLARRAEFAAGLRQOKR 394

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP-archaea:*
2: SP-bacteria:*
3: SP-fungi:*
4: SP-human:*
5: SP-invertebrate:*
6: SP-mammal:*
7: SP-mhc:*
8: SP-organelle:*
9: SP-phage:*
10: SP-plant:*
11: SP-rodent:*
12: SP-virus:*
13: SP-vertebrate:*
14: SP-unclassified:*
15: SP-virus:*
16: SP-bacteriap:*
17: SP-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	994	50.1	397	16	08UC86
2	754	38.0	472	10	09SP64
3	747	37.7	473	10	09FTN0
4	746	37.6	472	10	09FX27
5	745	37.6	475	10	0947C3
6	740	37.3	473	10	08W250
7	715	36.0	474	10	09M4W4
8	684.5	34.5	382	16	08RA28
9	665.5	33.5	386	2	09AUD7
10	568	28.6	390	16	08R622
11	519.5	26.2	488	5	09G693
12	181.5	9.1	115	2	09EX11
13	136.5	6.9	2555	11	091ZB6
14	136.5	6.9	2561	11	08VIF5
15	133.5	6.7	412	16	09A4V4
16	130.5	6.6	981	16	08XS09

17	128.5	6.5	442	16	098FT2	098ft2 rhizobium 1
18	124	6.2	439	16	08XZ88	08xzt8 ralstonia s
19	123.5	6.2	1731	16	08U8W4	08u8w4 agrobacteri
20	122.5	6.2	3247	12	065553	065553 bovine herp
21	121.5	6.1	585	16	09ADP2	09adp2 streptomyc
22	121	6.1	1041	2	09JC90	09jtc90 actinomadr
23	119.5	6.0	5435	2	09LX2	09lxc2 streptomyc
24	119	6.0	538	16	050442	050442 mycobacteri
25	118.5	6.0	4881	2	09SOR3	09sor3 streptomyc
26	118	5.9	437	2	09RAM9	09ram9 methylobaci
27	117	5.9	520	16	09A436	09a436 caulobacter
28	115.5	5.8	434	16	08YD51	08yds1 brucella me
29	115.5	5.8	9477	2	09LX3	09lxc3 streptomyc
30	115	5.8	1728	2	054591	054591 amycolatops
31	115	5.8	3190	2	093NW7	093nw7 streptomyc
32	114.5	5.8	3970	2	093H18	093h18 streptomyc
33	113.5	5.7	7257	2	09L8C7	09l8c7 polyanthum
34	113.5	5.7	7257	2	09K1Z7	09k1z7 polyanthum
35	112	5.6	951	2	068078	068078 rhodobacter
36	112	5.6	711	16	096855	096855 mycobacteri
37	112	5.6	1472	16	09L214	09l214 streptomyc
38	111.5	5.6	985	2	09ZNT6	09znt6 pseudomonas
39	111.5	5.6	1907	2	093H19	093h19 streptomyc
40	110.5	5.6	429	16	09RD83	09rd83 streptomyc
41	110.5	5.6	554	16	092T76	092t76 rhizobium m
42	110.5	5.6	941	16	092IWS	092iws streptomyc
43	110.5	5.6	1099	16	09A4D6	09a4d6 caulobacter
44	110.5	5.6	1281	16	09I2T1	09i2t1 pseudomonas
45	110.5	5.6	1293	2	005170	005170 streptomyc

ALIGNMENTS.

RESULT 1

ID	Q8UC86	PRELIMINARY:	PRT:	397 AA.
AC	08UC86;			
DT	01-JUN-2002 (TREMBLrel. 21, Created)			
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	1-deoxy-D-xylulose 5-phosphate reductoisomerase.			
GN	DXR OR ATU2617 OR AGR_C_4736.			
OS	Agrobacterium tumefaciens (strain C58 / ATCC 33970).			
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;			
OC	Rhizobiaceae; Rhizobium.			
OX	NCBI_TaxID=176299;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21608550; PubMed=11743193;			
RA	Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,			
RA	Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,			
RA	Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,			
RA	Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,			
RA	Kutyavin T., Levy R., Li M.-J., McCelland E., Palmeri A.,			
RA	Raymond C., Rouse G., Seemphimachak C., Wu Z., Romero P., Gordon D.,			
RA	Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,			
RA	Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,			
RA	Chunley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,			
RA	Nester E.W.;			
RT	"The genome of the natural genetic engineer Agrobacterium tumefaciens			
RT	C58.";			
RL	Science 294:2317-2323(2001).			
RL	(2)			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21608551; PubMed=11743194;			
RA	Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,			
RA	Quorillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,			
RA	Hommel K., Gordon J., Vaudin M., Iarchouk O., Epp A., Liu F.,			
RA	Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,			
RA	Flanagan C., Crowell C., Guron J., Lomo C., Sear C., Strub G.,			
RA	Cielo C., Slater S.;			
RT	"Genome sequence of the plant pathogen and biotechnology agent			

RT Agrobacterium tumefaciens C58.
 RL Science 294:2323-2328(2001).
 DR EMBL: AE009208; AAL43593.1;
 DR EMBL: AE008174; AAK8334.1;
 KW Isomerase; Complete proteome.
 SO SEQUENCE 397 AA; 42303 MW; 0A24C408C2D47094 CRC64;

Query Match 50.1%; Score 994; DB 16; Length 397;
 Best Local Similarity 50.4%; Pred. No. 1.3e-54;
 Matches 197; Conservative 64; Mismatches 126; Indels 4; Gaps 1;

QY 2 RSLIFGATGSGISTFDLVNRKGGPEAFRTVALTGGRNIRRLAEMARALKAEIAYTAHE 61
 Db KRLTILGSGISGNTLDVNRVROGSGDEFEIEMALGAGNIALLAQAEFGAQLAVTADD 68
 QY 62 DCLPALREALGTEVAGCAIAEAADRPADWTMSAIVAGLVPGMRALKRGRTAL 121
 Db 69 DKYEALKSALAGTGKVAAGAAAGLEAASMDAGVMMAIACTPGLAPTLTARACADIAL 128
 QY 122 ANKESLYTAGOLMRTAOENGATLLPVDESHSAVFOALAGEDTACVERVITTAAGCP 181
 Db 129 ANKEELVAGDVFRTYVGGGRLIPVDESHSAIFOCITGEYKQAVERTIVLTASGCPERT 188
 QY 182 WSLERIRACTVAEAOAHNMSMGORISIDSAMFNKALELETRFEFGEDRIEAVVHP 241
 Db 189 WSRDEMSVNTADIAHAPNMSGLKVSIGSASMFNKGLEMEATKYLFEDLRPDQDVIVHP 248
 QY 242 QSIYHAWVFCDDGIMAHILGPADMRHAIGFALNMPGREGVAVARIDLAOIASLTFOKPE 301
 Db 249 QSIHSMVGYTDGSIYAGLGSPDKRTATSVALTYPGERNLVERLDFAKLRLDPEARDE 308
 QY 302 ERFPALRLARVMAARGLSGAFAFNKAEIALDHFIAGTIGFLDMAAVEELVAGVSTDP 361
 Db 309 ARFPALRLARALRRGGIGGALNAEETAFAHFAVGIGFLDMAEIVTMDRMDH--- 365
 QY 362 FGKVPDALEEVYAMDHLARRAEFAAGLRQ 392
 Db 366 -GRTAETMDVFSADDEARRALIELATKEK 395

RESULT 2

Q9SP64 PRELIMINARY; PRT; 472 AA.
 AC Q9SP64; 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE 1-deoxy-D-xylose-5-phosphate reductoisomerase.
 GN DKRL.
 OS Artemisia annua (Sweet wormwood).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
 OC Anthemideae; Artemisia.
 OC NCBI_TaxID=35608;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. YUGOSLAVIA; TISSUE=ROOT;
 RA Mobbie K.K., Souret F.F., Shore K.A., Weathers P.J.;
 RT "Artemisia annua 1-deoxy-D-xylose-5-phosphate reductoisomerase (dxr)
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF182287; AAD56391.2;
 DR InterPro: IPR003821; DXP_reductisomase.
 DR Pfam: PF02670; DXP_reductoisom; 1.
 DR TIGRFAMs: TIGR00243; Dxr; 1.
 KW Isomerase.
 SO SEQUENCE 472 AA; 50740 MW; D52023C09D475675 CRC64;

Query Match 38.0%; Score 754; DB 10; Length 472;
 Best Local Similarity 41.1%; Pred. No. 1.7e-39;
 Matches 164; Conservative 77; Mismatches 136; Indels 22; Gaps 7;

QY 2 RSLIFGATGSGISTFDLVNRKGGPEAFRTVALTGGRNIRRLAEMARALKAEIAYTAHE 61
 Db KRLTILGSGISGNTLDVNRVROGSGDEFEIEMALGAGNIALLAQAEFGAQLAVTADD 68
 QY 62 DCLPALREALGTEVAGCAIAEAADRPADWTMSAIVAGLVPGMRALKRGRTAL 118
 Db 69 DKYEALKSALAGTGKVAAGAAAGLEAASMDAGVMMAIACTPGLAPTLTARACADIAL 128
 QY 122 ANKESLYTAGOLMRTAOENGATLLPVDESHSAVFOALAGEDTACVERVITTAAGCP 181
 Db 129 ANKEELVAGDVFRTYVGGGRLIPVDESHSAIFOCITGEYKQAVERTIVLTASGCPERT 188
 QY 182 WSLERIRACTVAEAOAHNMSMGORISIDSAMFNKALELETRFEFGEDRIEAVVHP 241
 Db 189 WSRDEMSVNTADIAHAPNMSGLKVSIGSASMFNKGLEMEATKYLFEDLRPDQDVIVHP 248
 QY 242 QSIYHAWVFCDDGIMAHILGPADMRHAIGFALNMPGREGVAVARIDLAOIASLTFOKPE 301
 Db 249 QSIHSMVGYTDGSIYAGLGSPDKRTATSVALTYPGERNLVERLDFAKLRLDPEARDE 308
 QY 302 ERFPALRLARVMAARGLSGAFAFNKAEIALDHFIAGTIGFLDMAAVEELVAGVSTDP 361
 Db 309 ARFPALRLARALRRGGIGGALNAEETAFAHFAVGIGFLDMAEIVTMDRMDH--- 365
 QY 362 FGKVPDALEEVYAMDHLARRAEFAAGLRQ 392
 Db 366 -GRTAETMDVFSADDEARRALIELATKEK 395

RESULT 3

Q9FTNO PRELIMINARY; PRT; 473 AA.
 AC Q9FTNO; 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE 1-deoxy-d-xylose-5-phosphate reductoisomerase.
 GN P0005A05.19 OR P0462C06.2.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OC NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponebare(GA3) genomic DNA, chromosome 1, PAC
 RL clone:P0005A05." to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponebare(GA3) genomic DNA, chromosome 1, PAC
 RL clone:P0482C06." to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP002863; BAB16915.1;
 DR EMBL: AP002845; BAB78606.1;
 DR InterPro: IPR003821; DXP_reductisomase.
 DR Pfam: PF02670; DXP_reductoisom; 1.
 DR TIGRFAMs: TIGR00243; Dxr; 1.
 KW Isomerase.
 SO SEQUENCE 473 AA; 51473 MW; 9933D9D8C3D1FC49 CRC64;

Query Match 37.7%; Score 747; DB 10; Length 473;
 Best Local Similarity 41.4%; Pred. No. 4.7e-39;
 Matches 165; Conservative 74; Mismatches 138; Indels 22; Gaps 7;

QY 2 RSLIFGATGSGISTFDLVNRKGGPEAFRTVALTGGRNIRRLAEMARALKAEIAYTAHE 61
 Db KRLTILGSGISGNTLDVNRVROGSGDEFEIEMALGAGNIALLAQAEFGAQLAVTADD 68
 QY 62 DCLPALREALGTEVAGCAIAEAADRPADWTMSAIVAGLVPGMRALKRGRTAL 118
 Db 69 DKYEALKSALAGTGKVAAGAAAGLEAASMDAGVMMAIACTPGLAPTLTARACADIAL 128
 QY 122 ANKESLYTAGOLMRTAOENGATLLPVDESHSAVFOALAGEDTACVERVITTAAGCP 181
 Db 129 ANKEELVAGDVFRTYVGGGRLIPVDESHSAIFOCITGEYKQAVERTIVLTASGCPERT 188
 QY 182 WSLERIRACTVAEAOAHNMSMGORISIDSAMFNKALELETRFEFGEDRIEAVVHP 241
 Db 189 WSRDEMSVNTADIAHAPNMSGLKVSIGSASMFNKGLEMEATKYLFEDLRPDQDVIVHP 248
 QY 242 QSIYHAWVFCDDGIMAHILGPADMRHAIGFALNMPGREGVAVARIDLAOIASLTFOKPE 301
 Db 249 QSIHSMVGYTDGSIYAGLGSPDKRTATSVALTYPGERNLVERLDFAKLRLDPEARDE 308
 QY 302 ERFPALRLARVMAARGLSGAFAFNKAEIALDHFIAGTIGFLDMAAVEELVAGVSTDP 361
 Db 309 ARFPALRLARALRRGGIGGALNAEETAFAHFAVGIGFLDMAEIVTMDRMDH--- 365
 QY 362 FGKVPDALEEVYAMDHLARRAEFAAGLRQ 392
 Db 366 -GRTAETMDVFSADDEARRALIELATKEK 395

[illegible]


```

Db 240 VHPDSVHSHVEYVDGSIIMOGVANKMPTILAFYPEKEYNSSINFLDIKNNLT 299
Oy 297 QKPEERPARLARDVVAARGLSGAFAFNKAETALDHFINGRIGELVMAVETLAC- 355
Db 300 EADRKVKLGIDLAVRAGRTDTPYTVNASELAVELMKQIKFLDIYRIIEAMNSH 359
Oy 356 ---VSTDPLFGKVPDALEVLAMDLARRAAE 385
Db 360 QVLTNTD-----NALNVIKEVDREIRKKVRE 386

RESULT 11
O96693 PRELIMINARY; PRT; 488 AA.
AC 096693;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE 1-deoxy-D-xylose 5-phosphate reductoisomerase.
GN DXR.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H3;
RA Altincsek B., Sanderbrand S., Wiesner J., Jomaa H.;
RT "dxr as a potential target for antimalarial drugs.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF111813, AAD03739.1;
DR InterPro: IPR003821; DXP_reductoisomase.
DR Pfam: PF02670; DXP_reductoisom; 1.
DR TrEMBL: TIGR00243; Dxr; 1.
KW isomerase.
SQ SEQUENCE 488 AA; 55756 MW; 4E280C81CDFAD3EF CRC64;

Query Match 26.2%; Score 519.5; DB 5; Length 488;
Best Local Similarity 30.5%; Pred. No. 8.5e-25;
Matches 124; Conservative 88; Mismatches 167; Indels 27; Gaps 8;

Oy 3 SLSIFGATSGISGSPFDLVMRKGPE-AFTVALTGSNIRLMAALAEAVTAE 61
Db 79 NVALIFGSGISGTMALNIRCKNIENFVNKALVNSVELLEQARELPYICLHDK 138
Oy 62 DCLPALREALGTGE--VAGGAQAIAE-AADRPADWTSAGIAGLVGMRALKGR 117
Db 139 SYBELKELVKNIKDYRIILCGDEGMKEICSSNSIDKIVIGISFQGLYSFMVAIMNK 198
Oy 118 TLALANKESLYTAGOLMRTAQ-ENGATILPVDSHSAVFPALAGE---DTAC----- 166
Db 199 IVALANKESIVSAGFEFLKLNIRKNAKIIPVDSHSAIFQCLDNKVKLTKCLQDNFSK 258
Oy 167 ---VERVITFASGSPFRMSLERIACVAAEQAHPNMSMORISIDSSAFENKALTELIE 223
Db 259 INNINKIFLCSGGPFQULMDELKNTSENAKHPKMKKTKITIDATMANKGLEVIE 318
Oy 224 TRFEGEPRIEAVVHQSIIVHAMVGCGLMAHLGAPADMRAHIGFALNMPGGEVAV 283
Db 319 THFLFDVVDYNDIEVIVHECIIHSCVEFIDKSVISQMYYPDMQIPILYSLWPDRIKTNL 378
Oy 284 ARIDLAQASTFPKPEERPARLARDVVAARGLSG---AENAKETALDHFINGR 339
Db 379 KPLDLAOSTLTTEHFKPSLEHPCIKLAAQA---GIKNFYPYVLAWSNEIANMLFLNKK 434
Oy 340 IGFLDMAVVEETLAGVSTDLFGKVPDALEVLAMDLARRAAE 385
Db 435 IKYDISSIIISQVLESFNSQVSENSEDLMOIILQIHSMAKDRAKD 480

RESULT 12
O9EX11 PRELIMINARY; PRT; 115 AA.
ID O9EX11

```

```

AC 09EX11;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE DXR reductoisomerase (Fragment).
GN DXR.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CG43;
RA Lai Y.C., Chang H.Y., Peng H.L.;
RT "Identification of Klebsiella pneumoniae CG43 genes specifically
expressed in vivo.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ292312; CAC19426.1;
DR InterPro: IPR003821; DXP_reductoisomase.
DR Pfam: PF02670; DXP_reductoisom; 1.
KW isomerase.
FT NON_TER
SQ SEQUENCE 115 AA; 12174 MW; 0B5E97D06F644C8C CRC64;

Query Match 9.1%; Score 181.5; DB 2; Length 115;
Best Local Similarity 41.9%; Pred. No. 0.0002;
Matches 49; Conservative 14; Mismatches 49; Indels 5; Gaps 3;

Oy 1 MRSISFGATSGISGSPFDLVMRKGPEAFPTVALTGSNIRLMAALAEAVTAE 60
Db 1 MKQLVLSGTSIGSTLDVY--RHNGRFSVALVACKNVDRAVEQCELTTPRYAVMD 58

Oy 61 EDCJPALREAL-AGTGTAVAGGAQAIAE-AADRPADWTSAGIAGLVGMRALK 114
Db 59 AQSAERLRLTRHSGSRFVLSGQAAAEVAALDEVDQVMAIYGAAGLVPTLAIR 115

RESULT 13
O91Z6 PRELIMINARY; PRT; 2555 AA.
ID O91Z6
AC 091Z6;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Beta4-spectrin.
GN SPNB4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21419653; PubMed=11528393;
RA Parkison N.J., Olsson C.L., Hallows J.L., McKee-Johnson J.,
RA Keogh B.P., Noben-Trauth K., Kujawa S.G., Tempel B.L.;
RT "Mutant beta-spectrin 4 causes auditory and motor neuropathies in
quivering mice.";
RL Nat. Genet. 29:61-65(2001).
DR EMBL: AY032655; AAK38731.1;
DR MGD: MGI:1890574; Spnb4.
DR InterPro: IPR001589; Actbind_actinin.
DR InterPro: IPR001715; Calponin-like.
DR InterPro: IPR001849; PH.
DR InterPro: IPR002017; Spectrin.
DR Pfam: PF00307; CH; 2.
DR Pfam: PF00169; PH; 1.
DR Pfam: PF00435; spectrin; 17.
DR PROSITE: PS00019; ACTININ_1; UNKNOWN_1.
DR PROSITE: PS00020; ACTININ_2; UNKNOWN_1.
DR PROSITE: PS50021; CH; 2.
DR PROSITE: PS50003; PH DOMAIN; 1.
SQ SEQUENCE 2555 AA; 288132 MW; 39FBCTE410289D45 CRC64;

```


Query Match 6.7%: Score 133.5; DB 16; Length 412;
Best Local Similarity 24.9%; Pred. No. 0.99;
Matches 98; Conservative 43; Mismatches 161; Indels 91; Gaps 19;

Db 89 SNAETTFSLGCRLELGGKR--RTVAATAQAEVETAARAAIRADLTQEY-----KTRY 140
D 12 SIGSTEDLVNR--KGGEAFPTVALTGGRNIRILAEMARKALAEAVTAHEDCLPALRE 69
OY 70 ALAGGTGVACGAQAIAARDPADWT-MSAIVGAAGLVPGRRALKHGRTTLANKESLY 128
DB 141 ALA---IDAKRVAVAQDAERAPRLVAALVVGSGEPPRLRAR-ART-ASDEAEAAV 195

Db 805 RRLAQHQHNLAGE--VEAHRCFVGGLRKQLATLGASGAGPLVVALQYRVAEADLF AEV 862
OY 87 -EAADRPAWMTSAIVCGAAGLVPGRRALKHGRTTLANKESLYVTAGOLMPTAOENGATI 145
Db 863 TEVALLRQWLRLDALA-----YYRRFGVHACELMIIGEQL-----LAMR----- 904
OY 146 LPVDSEHSAY----FOALAGEDPTACVERI-----ITASGGFFRWMLSERIRACTYA 193
Db 905 VPDSLDDEVVQYHRPESLDQEMNSLMGRVLDTNQTVQELVEGHP-----SSDEVSC--- 957
OY 194 EQAAPPNMSMGORISIDSASMKNKLDELTEFEFGFPDRLEAVY----HQSI--VHA 247
Db 958 --QDILN-----SRWNRIVELVEORK-----EEMSALLVENNHLEVAEKRA 997
OY 248 MYGFCDGLMAHLGPDAMRHAI-----GFALNMPGRG-EYVARIDILOJASTFQRPD 300
Db 998 QY-----REKRRAVESAPRAGGALOMVLSGLEALQALPERQOALLEBAL 1044
OY 301 EERP--ALRL--ARDYMAARG-TSQAFAFNAKEIA---LDHFIAIGFLDMAAAYE 350
Db 1045 AERFPAQATRLHQGEELGEAGWALAGAQAQAGEVAAAGRLOREFRLDPDLMDLVRAQ 1104
OY 351 ETLAGVSTDPLEGKVPDALEVLANDHARRA-EAAGLRQ 392
Db 1105 EAAGAVE-----GPLRSLEEADGL-LARHALKEEVDOREE 1140

RESULT 15

ID 09A4V4 PRELIMINARY; PRY. 412 AA.
ID 09A4V4
AC 09A4V4:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Metal ion efflux outer membrane factor protein family.
GN CC2721.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
CC Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nieman W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Plosocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeJoy R.T., Dodson R.J., Durkin A.S., Gailin M.L., Halt D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Ueberlack T., Tran K., Wolf A., Yamnethvan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005938; AAK24686.1; .
DR TIGR; CC2721; .
DR InterPro; IPR003423; OEP.
DR Pfam; PF02321; OEP. 2.
KW Complete proteome.
SQ SEQUENCE 412 AA; 43076 MW; 53083C2167D0E7E6 CRC64;

```
QY 129 TAGOLMRTAENGATILPVDESHSAVFOALAGEDTACVERVIITASGPFPRWSLERIR 188
Db 196 LAAOTL-----ALA-----ARRALTSIMGDP--DFEIELVE 224
QY 189 --ACTVAEAOAHPPNWSKQORISIDSASMFNKALEIETREFFGFEPD-RIEAVHPOSIV 245
Db 225 PLAAPTQMAVDPNISTLDVRLAQAERAAAEAAIE---RERAGDPDLTIOA----- 272
QY 246 HAMVFCDDGIMAHLPADMHRAIGFALNMP---GRGEVVARID-----LAQIA 292
Db 273 -----GYRREFQOTDRALIVGFTAPIPIRDRNQCNVAAARADANAQAARERLALTG 323
QY 293 SLTFOKPEERFPALRLARDVMAARGLSGAAPNNAKETALDHTAGRIGFLDMAAVEET 352
Db 324 SIRAMRDAQASIKAAEAEALVLSSTVPOA--OOAVDLAROGFEAGKFSLLDVLDQAAL 381
QY 353 LAGVSTDPLEFGKVPDPALEEVLMMDHLARRAAE 385
Db 382 LA--SRNDL---IAARLERALALALERAAOE 409
```

Search completed: April 15, 2003, 13:59:51
Job time : 93 secs


```

Db 119 KTVLLANKKEALVMSGQIMQAVSDSGAVLLPIDSSEHNAIFQCMRGYTPGHAKQARRIL 178
Oy 172 ITASGCPFRDMSLEIRACTYAEAOAHPNMSGQRISIDSASMFNKALELLETRFPGFE 231
Db 179 LTASGCPFRPRPIETLSSVTPOAAVHPKMDGKRISVDSATMNMKGLELLEACLLFNME 238
Oy 232 PDRIEAVVHPOSIVAHVAFCDGGLMAHLGPADMRHAIGFALNMPGRCGEVVARIDLAOI 291
Db 239 PDQIEVVIHPOSIIHSMVDYDGSVLAQMGNPDMTPTLAHAAWPERDGSVAPLDIFEV 298
Oy 292 ASLTFQKDEERFPALRLAROVMAARGLSGAAFNAKEIALDHFYAGRIIGFLDMAAYVEE 351
Db 299 GHMDEKPDRLKRFPCRLAYELIKSGGIMPTVLNANAEIAYEAFINEEVKFTDIAVIER 358
Oy 352 TLGAVSTDPLEKGVDALEEVLAAMDHLARRAAEE 385
Db 359 SMAQFKPDD-----AGSLELVLAQDQDAREVARD 387

```

RESULT 2

```

US-09-934-903-4
: Sequence 4, Application US/09934903
: Patent No. US20020102690A1
: GENERAL INFORMATION:
: APPLICANT: Koffas, Mattheos
: APPLICANT: Odom, J. Martin
: APPLICANT: Schenzle, Andreas J.
: APPLICANT: No. US20020102690A1ton, Kelley C.
: APPLICANT: Tomb, Jean-Francois
: APPLICANT: Rouviere, Pierre
: APPLICANT: Picataggio, Stephen
: APPLICANT: Cheng, Qiong
: TITLE OF INVENTION: Genes Involved in Isoprenoid Compounds Production
: FILE REFERENCE: C11646 US NA
: CURRENT APPLICATION NUMBER: US/09/934,903
: PRIOR FILING DATE: 2001-08-22
: PRIOR APPLICATION NUMBER: 60/229,907
: NUMBER OF SEQ ID NOS: 24
: SOFTWARE: Microsoft Office 97
: SEQ ID NO 4
: LENGTH: 394
: TYPE: PR
: ORGANISM: Methylobionas 16a
: FEATURE:
: OTHER INFORMATION: Amino acid sequences encoded by ORF2
US-09-934-903-4

```

Query Match 39.2%; Score 778; DB 10; Length 394;

Best Local Similarity 41.6%; Pred. No. 5e-60; Matches 164; Conservative 75; Mismatches 139; Indels 16; Gaps 5;

```

Oy 1 MRSLSIFGATSGISTESTFDLVMRKGGPEAFRTVALTGGRNIRLAEMARALKAEIAYTAH 60
Db 1 MKGICILGATSGISGVSTLDVVARH--PDKYQVVALTANGNDALYEOCLAHHPREYAVVM 58
Oy 61 EDCIPALREALAGT--GTEVAGGAQAIAEAAD--RPADMTSAIYGAAGLVGRALKHG 116
Db 59 ESKAEEFKQRIASPVADIKVLSSGEALQOYATLENDVTVAALYGAAGLLPTLAANAAG 118
Oy 117 RTLLANKESLVTAAGOLMRTAENGATILPVDSHSAVFOAL-----AGEDTACVERVI 171
Db 119 KTVLLANKKEALVMSGQIFMAVSDSGAVLLPIDSSEHNAIFQCMRGYTPGHAKQARRIL 178
Oy 172 ITASGCPFRDMSLEIRACTYAEAOAHPNMSGQRISIDSASMFNKALELLETRFPGFE 231
Db 179 LTASGCPFRPRPIETLSSVTPOAAVHPKMDGKRISVDSATMNMKGLELLEACLLFNME 238
Oy 232 PDRIEAVVHPOSIVAHVAFCDGGLMAHLGPADMRHAIGFALNMPGRCGEVVARIDLAOI 291
Db 239 PDQIEVVIHPOSIIHSMVDYDGSVLAQMGNPDMTPTLAHAAWPERDGSVAPLDIFEV 298

```

```

Oy 292 ASLTFQKDEERFPALRLAROVMAARGLSGAAFNAKEIALDHFYAGRIIGFLDMAAYVEE 351
Db 299 GHMDEKPDRLKRFPCRLAYELIKSGGIMPTVLNANAEIAYEAFINEEVKFTDIAVIER 358
Oy 352 TLGAVSTDPLEKGVDALEEVLAAMDHLARRAAEE 385
Db 359 SMAQFKPDD-----AGSLELVLAQDQDAREVARD 387

```

RESULT 3

```

US-09-934-868-64
: Sequence 64, Application US/09934868
: Patent No. US20020137190A1
: GENERAL INFORMATION:
: APPLICANT: Koffas, Mattheos
: APPLICANT: Odom, James M
: APPLICANT: Schenzle, Andreas J
: TITLE OF INVENTION: DENTRIFYING METHANOTROPHIC BACTERIAL STRAIN
: FILE REFERENCE: C11596 US NA
: CURRENT APPLICATION NUMBER: US/09/934,868
: PRIOR FILING DATE: 2001-08-22
: PRIOR APPLICATION NUMBER: 60/229,858
: NUMBER OF SEQ ID NOS: 81
: SOFTWARE: Microsoft Office 97
: SEQ ID NO 64
: LENGTH: 394
: TYPE: PR
: ORGANISM: Methylobionas 16a
: FEATURE:
: OTHER INFORMATION: Amino acid sequences encoded by DXR
US-09-934-868-64

```

Query Match 39.2%; Score 778; DB 10; Length 394;

Best Local Similarity 41.6%; Pred. No. 5e-60; Matches 164; Conservative 75; Mismatches 139; Indels 16; Gaps 5;

```

Oy 1 MRSLSIFGATSGISTESTFDLVMRKGGPEAFRTVALTGGRNIRLAEMARALKAEIAYTAH 60
Db 1 MKGICILGATSGISGVSTLDVVARH--PDKYQVVALTANGNDALYEOCLAHHPREYAVVM 58
Oy 61 EDCIPALREALAGT--GTEVAGGAQAIAEAAD--RPADMTSAIYGAAGLVGRALKHG 116
Db 59 ESKAEEFKQRIASPVADIKVLSSGEALQOYATLENDVTVAALYGAAGLLPTLAANAAG 118
Oy 117 RTLLANKESLVTAAGOLMRTAENGATILPVDSHSAVFOAL-----AGEDTACVERVI 171
Db 119 KTVLLANKKEALVMSGQIFMAVSDSGAVLLPIDSSEHNAIFQCMRGYTPGHAKQARRIL 178
Oy 172 ITASGCPFRDMSLEIRACTYAEAOAHPNMSGQRISIDSASMFNKALELLETRFPGFE 231
Db 179 LTASGCPFRPRPIETLSSVTPOAAVHPKMDGKRISVDSATMNMKGLELLEACLLFNME 238
Oy 232 PDRIEAVVHPOSIVAHVAFCDGGLMAHLGPADMRHAIGFALNMPGRCGEVVARIDLAOI 291
Db 239 PDQIEVVIHPOSIIHSMVDYDGSVLAQMGNPDMTPTLAHAAWPERDGSVAPLDIFEV 298
Oy 292 ASLTFQKDEERFPALRLAROVMAARGLSGAAFNAKEIALDHFYAGRIIGFLDMAAYVEE 351
Db 299 GHMDEKPDRLKRFPCRLAYELIKSGGIMPTVLNANAEIAYEAFINEEVKFTDIAVIER 358
Oy 352 TLGAVSTDPLEKGVDALEEVLAAMDHLARRAAEE 385
Db 359 SMAQFKPDD-----AGSLELVLAQDQDAREVARD 387

```

RESULT 4

```

US-09-923-556-6
: Sequence 6, Application US/09923556
: Patent No. US20020058321A1
: GENERAL INFORMATION:
: APPLICANT: Busch, Marco
: APPLICANT: Hain, Rudiger

```

```

; APPLICANT: Martin, William
; APPLICANT: Tietjen, Klaus
; APPLICANT: Klotz, Andreas
; TITLE OF INVENTION: Method of determining the activity of
; TITLE OF INVENTION: 1-deoxy-D-xylose-5-phosphate reductoisomerase and
; TITLE OF INVENTION: 1-deoxy-D-xylose-5-phosphate synthase
; FILE REFERENCE: 2020US
; CURRENT APPLICATION NUMBER: US/09/923,556
; PRIOR FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/449,335
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 6
; LENGTH: 477
; TYPE: PRF
; ORGANISM: Arabidopsis thaliana
US-09-923-556-6
```

```

Query Match          37.4%; Score 741.5; DB 10; Length 477;
Best Local Similarity 41.3%; Pred. No. 9,9e-57;
Matches 162; Conservative 71; Mismatches 140; Indels 19; Gaps 6;
```

```

QY 2 RSLSTFGATGSGESTFDLMRRKGPPAFTVALTGGNTRRLAEMARALKAELAVTAHE 61
   :|||:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 81 KPISIVGSGTSGTQTLDIYAE--NPDKFRVALAAGSNVTLADQVRRKPPALVAVRNE 138
   :|||:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 62 DCLPALREALAGTG--TEVAGGAOATAEADRP-ADMTSAIYGAAGLVGMRALKHRT 118
   :|||:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 139 SLINELKEALADLDYKLEITIPGEGVLEVARHPEAVTVGVIGACALKFTVAITEGKD 198
   :|||:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 119 LALANKESVLTAGQLMRTAENGATILPYDSEHSAVFOALAGEDTACVERVIITASGCP 178
   :|||:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 199 IALANKETLLAGGPFVPLANKHNKVLTPADSEHSAIFQCIQGLPBGALRKIIITASGA 258
   :|||:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 179 FRWMSLEIRACTYAEAOAHPPNMSGCRISIDSASFNKALELIEFFGFEPDRLEAV 238
   :|||:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 259 FRWMPVEKLEKVVADALAKHPNMNGKKITVDSATLFNKGLIEVAHYLFGAEYDDIEIV 318
   :|||:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 239 VHPQSTIVHAWGFCGGLMAHLGPDAMRHAIGFALNMPGR---GEVVARIDLAQIASLT 295
   :|||:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 319 IHQSTIHSMTETQDSSVLAQLGMPDKRLPLITMSWPDVPCSEVWPRDLCKLSLT 378
   :|||:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 296 FOKPDERFPALRLARDVMAARGLSGAAPNAKEIALDHFIAGRIGFLDMAAVEETL-- 353
   :|||:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 379 FKRPDVKYPSMDLAYAAGAGGTMTGVLSAANEKAVEMFIDEKISYLDIFKVELCDK 438
   :|||:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 354 --AGVSTDPLEKVPDALEEVLAHDLARRAA 383
   :|||:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 439 HRNELVTS-----SLEIIVHYDLMAREYA 463
   :|||:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
```

```

RESULT 5
; Sequence 2, Application US/09987025
; Patent No. US20020108148A1
; GENERAL INFORMATION:
; APPLICANT: Boronat, Albert
; APPLICANT: Campos, Narcisco
; APPLICANT: Kishore, Ganesh M.
; TITLE OF INVENTION: Nucleic Acid Sequences Involved in
; TITLE OF INVENTION: Isoprenoid Synthesis
; FILE REFERENCE: 17142/02/US
; CURRENT APPLICATION NUMBER: US/09/987,025
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 09/549,787
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/129,899
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/146,461
; PRIOR FILING DATE: 1999-07-30
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
```

```

; SEQ ID NO 2
; LENGTH: 477
; TYPE: PRF
; ORGANISM: Arabidopsis sp
US-09-987-025-2
```

```

Query Match          37.4%; Score 741.5; DB 10; Length 477;
Best Local Similarity 41.3%; Pred. No. 9,9e-57;
Matches 162; Conservative 71; Mismatches 140; Indels 19; Gaps 6;
```

```

QY 2 RSLSTFGATGSGESTFDLMRRKGPPAFTVALTGGNTRRLAEMARALKAELAVTAHE 61
   :|||:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 81 KPISIVGSGTSGTQTLDIYAE--NPDKFRVALAAGSNVTLADQVRRKPPALVAVRNE 138
   :|||:|||||:|:~|:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
QY 62 DCLPALREALAGTG--TEVAGGAOATAEADRP-ADMTSAIYGAAGLVGMRALKHRT 118
   :|||:|||||:|:~|:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
Db 139 SLINELKEALADLDYKLEITIPGEGVLEVARHPEAVTVGVIGACALKFTVAITEGKD 198
   :|||:|||||:|:~|:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
QY 119 LALANKESVLTAGQLMRTAENGATILPYDSEHSAVFOALAGEDTACVERVIITASGCP 178
   :|||:|||||:|:~|:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
Db 199 IALANKETLLAGGPFVPLANKHNKVLTPADSEHSAIFQCIQGLPBGALRKIIITASGA 258
   :|||:|||||:|:~|:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
QY 179 FRWMSLEIRACTYAEAOAHPPNMSGCRISIDSASFNKALELIEFFGFEPDRLEAV 238
   :|||:|||||:|:~|:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
Db 259 FRWMPVEKLEKVVADALAKHPNMNGKKITVDSATLFNKGLIEVAHYLFGAEYDDIEIV 318
   :|||:|||||:|:~|:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
QY 239 VHPQSTIVHAWGFCGGLMAHLGPDAMRHAIGFALNMPGR---GEVVARIDLAQIASLT 295
   :|||:|||||:|:~|:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
Db 319 IHQSTIHSMTETQDSSVLAQLGMPDKRLPLITMSWPDVPCSEVWPRDLCKLSLT 378
   :|||:|||||:|:~|:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
QY 296 FOKPDERFPALRLARDVMAARGLSGAAPNAKEIALDHFIAGRIGFLDMAAVEETL-- 353
   :|||:|||||:|:~|:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
Db 379 FKRPDVKYPSMDLAYAAGAGGTMTGVLSAANEKAVEMFIDEKISYLDIFKVELCDK 438
   :|||:|||||:|:~|:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
QY 354 --AGVSTDPLEKVPDALEEVLAHDLARRAA 383
   :|||:|||||:|:~|:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
Db 439 HRNELVTS-----SLEIIVHYDLMAREYA 463
   :|||:|||||:|:~|:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
```

```

RESULT 6
; Sequence 29, Application US/10047412A
; Publication No. US20020197696A1
; GENERAL INFORMATION:
; APPLICANT: Levin, Joshua Z.
; APPLICANT: Budziszewski, Gregory J.
; APPLICANT: Potter, Sharon L.
; APPLICANT: Weigich, Lynette M.
; TITLE OF INVENTION: Herbicide Target Genes and Methods
; FILE REFERENCE: PB/5-30780DIV
; CURRENT APPLICATION NUMBER: US/10/047,412A
; CURRENT FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 29
; LENGTH: 477
; TYPE: PRF
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (39)
; OTHER INFORMATION: Xaa = Asp or Asn
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (155)
; OTHER INFORMATION: Xaa = Pro or Leu
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (162)
; OTHER INFORMATION: Xaa = Leu or Glu
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (187)
```

OTHER INFORMATION: Xaa - Lys or Glu
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (465)
 OTHER INFORMATION: Xaa - Asp or Asn
 US-10-047-412A-29

Query Match 37.2%; Score 737.5; DB 9; Length 477;
 Best Local Similarity 41.1%; Pred. No. 2.2e-56;
 Matches 161; Conservative 71; Mismatches 141; Indels 19; Gaps 6;

QY 2 RSLIFGATGSGESTFDLVNRKGGPEAFRTVALTGGRNIRRLAEMARALAEAVTAHE 61
 DB 81 KPISTVSTGSGIGTDLIDVAE--NPDKFRVVALAAGSNVTLADQVRRFPALVAAVNE 138
 QY 62 DCLPALREALAGTG--TEVAGGAQAIAEAADRP--ADMTMSAIVGAGLVPGMRALKHGRT 118
 DB 139 SLINELKALADLDYKKEIIIGEXGVIEVARHPREAVTVTIGCGAGLKPVAIAEAGKD 198
 QY 119 LALANKESLVTAGOLMNTAENGATILPVDESHSAVQALAGEDTACVERVIITASGCP 178
 DB 199 IALANKETLLIAGGPFVPLANKHNKILPADSEHSAIFQCIQGLPEGALRKIIILITASGGA 258
 QY 179 FRDMSLERIRACTVAEAOAHNMGSGORISIDSAMFNKALELLETREFFGEPDRIEAV 238
 DB 259 FRDMPVEKLEKVVADALKHNNMNGKITYDSCITLFNKGLEVEIAHLYFGAEYDDIEIV 318
 QY 239 VHPOSIYAMVGFCDGGLMAHLGPADMRAHATGALMMPGR--GEVPARIDLAQIASLT 295
 DB 319 IHPOSIHSMIETODSSVLAQLGWPDMLPLITYMSWPDRCSEVTPRDLCKLSLT 378
 QY 296 FOKPDEERFPLRLARDVMAARGLSGAAFNAKELADHFTAGRIGFLDMAAAYVEETL-- 353
 DB 379 FKRPDNNKYPSPMDLAAYAAGRGAGTMTGVLSAANEKAVEMFIDEKISTYLDIKFVVELTCDK 438
 QY 354 --AGVSTDPFLGKVPDALEEVLAADHLARRAA 383
 DB 439 HRNELVTS-----SLEIIVHYDLMAREYA 463

RESULT 7
 US-09-923-556-2
 Sequence 2, Application US/09923556
 Patent No. US20020058321A1
 GENERAL INFORMATION:
 APPLICANT: Busch, Marco
 APPLICANT: Hain, Rudiger
 APPLICANT: Martin, William
 APPLICANT: Tietjen, Klaus
 APPLICANT: Klotz, Andreas
 TITLE OF INVENTION: Method of determining the activity of
 TITLE OF INVENTION: 1-deoxy-D-xylulose-5-phosphate reductoisomerase and
 FILE REFERENCE: 2020US
 CURRENT APPLICATION NUMBER: US/09/923,556
 PRIOR APPLICATION NUMBER: 2001-08-07
 PRIOR FILING DATE: EARLIER FILING DATE: 1999-11-24
 NUMBER OF SEQ ID NOS: 8
 SOFTWARE: Patent Ver. 2.1
 SEQ ID NO 2
 LENGTH: 477
 TYPE: PRT
 ORGANISM: Arabidopsis thaliana
 US-09-923-556-2

Query Match 37.2%; Score 737.5; DB 10; Length 477;
 Best Local Similarity 41.1%; Pred. No. 2.2e-56;
 Matches 161; Conservative 71; Mismatches 141; Indels 19; Gaps 6;

QY 62 DCLPALREALAGTG--TEVAGGAQAIAEAADRP--ADMTMSAIVGAGLVPGMRALKHGRT 118
 DB 139 SLINELKALADLDYKKEIIIGEXGVIEVARHPREAVTVTIGCGAGLKPVAIAEAGKD 198
 QY 119 LALANKESLVTAGOLMNTAENGATILPVDESHSAVQALAGEDTACVERVIITASGCP 178
 DB 199 IALANKETLLIAGGPFVPLANKHNKILPADSEHSAIFQCIQGLPEGALRKIIILITASGGA 258
 QY 179 FRDMSLERIRACTVAEAOAHNMGSGORISIDSAMFNKALELLETREFFGEPDRIEAV 238
 DB 259 FRDMPVEKLEKVVADALKHNNMNGKITYDSCITLFNKGLEVEIAHLYFGAEYDDIEIV 318
 QY 239 VHPOSIYAMVGFCDGGLMAHLGPADMRAHATGALMMPGR--GEVPARIDLAQIASLT 295
 DB 319 IHPOSIHSMIETODSSVLAQLGWPDMLPLITYMSWPDRCSEVTPRDLCKLSLT 378
 QY 296 FOKPDEERFPLRLARDVMAARGLSGAAFNAKELADHFTAGRIGFLDMAAAYVEETL-- 353
 DB 379 FKRPDNNKYPSPMDLAAYAAGRGAGTMTGVLSAANEKAVEMFIDEKISTYLDIKFVVELTCDK 438
 QY 354 --AGVSTDPFLGKVPDALEEVLAADHLARRAA 383
 DB 439 HRNELVTS-----SLEIIVHYDLMAREYA 463

RESULT 8
 US-10-047-412A-10
 Sequence 10, Application US/10047412A
 Publication No. US20020197696A1
 GENERAL INFORMATION:
 APPLICANT: Levin, Joshua Z.
 APPLICANT: Budziszewski, Gregory J.
 APPLICANT: Potler, Sharon L.
 APPLICANT: Megrich, Lynette M.
 TITLE OF INVENTION: Herbicide Target Genes and Methods
 FILE REFERENCE: PB/5-30780DIV
 CURRENT APPLICATION NUMBER: US/10/047,412A
 CURRENT FILING DATE: 2002-04-11
 NUMBER OF SEQ ID NOS: 29
 SOFTWARE: Patent Ver. 2.1
 SEQ ID NO 10
 LENGTH: 477
 TYPE: PRT
 ORGANISM: Arabidopsis thaliana
 US-10-047-412A-10

Query Match 37.1%; Score 736; DB 9; Length 477;
 Best Local Similarity 40.5%; Pred. No. 3e-56;
 Matches 163; Conservative 74; Mismatches 143; Indels 22; Gaps 7;

QY 2 RSLIFGATGSGESTFDLVNRKGGPEAFRTVALTGGRNIRRLAEMARALAEAVTAHE 61
 DB 81 KPISTVSTGSGIGTDLIDVAE--NPDKFRVVALAAGSNVTLADQVRRFPALVAAVNE 138
 QY 62 DCLPALREALAGTG--TEVAGGAQAIAEAADRP--ADMTMSAIVGAGLVPGMRALKHGRT 118
 DB 139 SLINELKALADLDYKKEIIIGEXGVIEVARHPREAVTVTIGCGAGLKPVAIAEAGKD 198
 QY 119 LALANKESLVTAGOLMNTAENGATILPVDESHSAVQALAGEDTACVERVIITASGCP 178
 DB 199 IALANKETLLIAGGPFVPLANKHNKILPADSEHSAIFQCIQGLPEGALRKIIILITASGGA 258
 QY 179 FRDMSLERIRACTVAEAOAHNMGSGORISIDSAMFNKALELLETREFFGEPDRIEAV 238
 DB 259 FRDMPVEKLEKVVADALKHNNMNGKITYDSCITLFNKGLEVEIAHLYFGAEYDDIEIV 318
 QY 239 VHPOSIYAMVGFCDGGLMAHLGPADMRAHATGALMMPGR--GEVPARIDLAQIASLT 295
 DB 319 IHPOSIHSMIETODSSVLAQLGWPDMLPLITYMSWPDRCSEVTPRDLCKLSLT 378
 QY 296 FOKPDEERFPLRLARDVMAARGLSGAAFNAKELADHFTAGRIGFLDMAAAYVEETL-- 353

Db 379 FKPDNVKYPMDLAYAGAGTMTGVLSPANAKAVEMIDEKISYLDIFKVELCDK 438
QY 354 --AGVSTDPLEFGKVPDALEEVLAHDLARRAAE---AAGLR 390
Db 439 HNELVTSP-----SLEIYHYDLMAREYADVOLSSGAR 473

RESULT 9

US-09-712-363-258
; Sequence 258 Application US/09712363
; Patent No. US20020164588A1
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, David
; APPLICANT: Rotstein, Sergio H.
; APPLICANT: Marcotte, Edward M.
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
; FILE REFERENCE: 07419-032001
; CURRENT FILING DATE: 2000-11-13
; PRIOR FILING DATE: 2000-01-28
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,531
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/117,844
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/118,206,
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/165,086
; NUMBER OF SEQ ID NOS: 1999-11-12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 258
; LENGTH: 436
; TYPE: PRF
; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-258

Query Match 31.7%; Score 628; DB 9; Length 436;

Best Local Similarity 39.2%; Pred. No. 6,9e-47;

Matches 155; Conservative 48; Mismatches 152; Indels 40; Gaps 7;

QY 6 IFGATGSGTSEFDFLVMRGGPEAFRTVALTGGRNIRRLAEMARALKAEIATVAHEDCLP 65
Db 40 VLASTOSTISTQALQVY--ADNPDREFVVGJLAA-----GAHLDTLL 78
QY 66 ALREALAGTGEVAG--GAQAI-----AEADR-----PADMTSAIYGAAGLVPGMR 111
Db 79 RQRAQGVNINAVADHAAQVRQDIPYHGSDATRLVEQTEADYVNLALVGALGRLPTLA 138
QY 112 ALKHGRTLLANKESLYTLAGOLIMRTAENGATILPYDSEHSAVFOALAGEPTACVERVY 171
Db 139 AKTGRLRLANKESLYVAGGSIVLRAARP--GQIVPDESHSLAQCIRGCTPDEYAKLV 196
QY 172 ITASGSPFDWMSLERIACVVAEAOAHPMWNGORISIDSASFENKALEIETREFEFGPE 231
Db 197 LTRASGPFPGWSAADLEHVTPEQAGHPWMSKGMNTLNSASLVNKGLEYIEHLLFGIP 256
QY 232 PRIEAVHPQSTIVHAMVFCDDGLMAHLGFPADMRHAIGFALNMPGREGVPAVRIOLAQI 291
Db 257 YDRIDVAVHPQSTIHSMTVFIDGSTIAQASPPMKLPISLALGMPRRVSGAAACPFHNA 316
QY 292 ASITPQKPEDEERFRLRLADYMAARGLSGAAPNAKETAIDHFIAGTIGFLDMAAVEEL 351
Db 317 SSWEFEPDLTDVFAVELARQAGVAGGCMTAVYVNAANEAAAFILAGRIGFPAIVGIAD 376

QY 352 TLAGVSTDPLEFGKVPDALEEVLAHDLARRAAEA 386
Db 377 VLHAADQ---WAVEPATVDVLDLAQRMARERARA 408

RESULT 10

US-09-738-626-5709
; Sequence 5709 Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAMA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO: 5709
; LENGTH: 392
; TYPE: PRF
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5709

Query Match 30.3%; Score 601; DB 9; Length 392;

Best Local Similarity 37.4%; Pred. No. 1.3e-44;

Matches 146; Conservative 67; Mismatches 155; Indels 22; Gaps 10;

QY 2 RLSITFGATGSGTSEFDFLVMRGGPEAFRTVAL--TGGNIRRLAEMAR--ALKAEIATVA 58
Db 6 KTLILGSTIGTQALDVI--ADNSDKREFVVGJLAA--GSDPLVISOAQLGLAARAV 63
QY 59 AHEDCLPALREALAG--TGTEVAGGAQAIAEADRPADMTSAIYGAAGLVPGMRALKH 115
Db 64 ADQAAAVASKALGELISTD---AAKILVETT--KADIVNALVAGSLGATLATLTLS 118
QY 116 GRTLLANKESLYTLAGOLIMRTAENGATILPYDSEHSAVFOALAGEPTACVERVYITTS 175
Db 119 GAHLALANKESLYVAGGEFV--TSKAKIQLIIPDSEHSAQAQIRGCTPDEYAKLVITTS 176
QY 176 GCFEPMWSLERIACVVAEAOAHPMWNGORISIDSASFENKALEIETREFEFGPEPDR 235
Db 177 GCFEPMWSLERIACVVAEAOAHPMWNGORISIDSASFENKALEIETREFEFGPEPDR 236
QY 236 EAVVHPQSTIVHAMVFCDDGLMAHLGFPADMRHAIGFALNMPGREGVPAVRIOLAQI 293
Db 237 DVTVHPQSTIHSMTVFIDGSTIAQASPPMKLPIALALDMPHR--VPKQAPLADPTAAHT 294
QY 294 LTRQKPEDEERFRLRLADYMAARGLSGAAPNAKETAIDHFIAGTIGFLDMAAVEEL 353
Db 295 WAFEPVDAAFPVAVOLARHAKOKGTPPAVYVNAANEAAEAFILAGRIGFPAIVGIAD 354
QY 354 AGVSTDPLEFGKVPDALEEVLAHDLARRAA 383
Db 355 QGASQ---FAGVASHVDILATESEARARA 381

RESULT 11

```

US-09-860-846-2
; Sequence 2, Application US/09860846
; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/860,846
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 5215
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-860-846-2

```

```

Query Match          5.8%; Score 115; DB 9; Length 5215;
Best Local Similarity 20.2%; Pred. No. 1.1;
Matches 108; Conservative 52; Mismatches 162; Indels 212; Gaps 21;

```

```

QY 28 EAFRTVALTGGRNIRRLAEAMARKAEIA-----VTANEDCLPALREALAGTGEVAG 80
DB 1509 QAVRLRTFLDGRPDVTVAADGRSLAARTAFEHKALTTATFDELLAGIDL-GRGEQATG 1567
QY 81 -----GAQAIA-----GLVP-----GMRALKH-GRITLANK 125
DB 1568 LVTGEPARAGRTAFLTGCGAQRVAMGEELRAHPVFAALDTVYALDRHLDRPLREIV 1627
QY 102 GAA-----GLVP-----GMRALKH-GRITLANK 125
DB 1528 AAGEELDTAVTOPALFAFEVALFRLEHGLVDPDLTGHSVGEIAAHVAGVSLDDAA 1687
QY 126 SLVTAGOLMRTAOENATILPVDSHSAVFOALAGEDTACVERVITTAGSGP----- 179
DB 1688 RLVTARGLMOSAREGA-MIAVAGAEVAVESLKGE-----GRVAAVAVGPTAVVYSG 1742
QY 180 -RDWSLERIRACTVAEAOAHNMS-MGQRI-----SIDASMFNKALELIERREFEGF 230
DB 1743 DAD-AAEIRAV-----MAGRGRTRLRVSHAFSPHMDVDLDEFRLVAEGITF 1791
QY 231 EPRDIEAVHPQSTVHAMV-----GFCDDGLMAHLGP 262
DB 1792 EEPRIPIV---STVTGALVTSGETLSPAYWVDQIRRPVRFDAVRTLAADATVLEIGP 1848
QY 263 ADMRAHAGFALNMPGRGEVPAVIDLAQIASLTQKPDDE----- 302
DB 1849 DAVITLAAEALAGT-DAPDAR-DVTIVPRLLRAGREPELTAAGLATAHVHGAPLDRAS 1906
QY 303 -----RPPALRLANDVMAARG--LSGA 322
DB 1907 FFPDGRRTDPTTAFRRHHWLTPEARTDARALGFDPAHRPLLTITVEVAGDGVLLTGR 1966
QY 323 AFNAKEIALDHFITAGRI-----GFLDMAVVEETLAGVSTDPLEFGVPDALEE 371
DB 1967 LSLTDQWMLADHMVNGAVLLPATATFELALAAAGDHVAGVAVVEELTLEAPLVLP 2020

```

```

RESULT 12
US-09-861-289-2
; Sequence 2, Application US/09861289
; Patent No. US20020110897A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin

```

```

; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/861,289
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 5215
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-861-289-2

```

```

Query Match          5.8%; Score 115; DB 10; Length 5215;
Best Local Similarity 20.2%; Pred. No. 1.1;
Matches 108; Conservative 52; Mismatches 162; Indels 212; Gaps 21;

```

```

QY 28 EAFRTVALTGGRNIRRLAEAMARKAEIA-----VTANEDCLPALREALAGTGEVAG 80
DB 1509 QAVRLRTFLDGRPDVTVAADGRSLAARTAFEHKALTTATFDELLAGIDL-GRGEQATG 1567
QY 81 -----GAQAIA-----GLVP-----GMRALKH-GRITLANK 125
DB 1568 LVTGEPARAGRTAFLTGCGAQRVAMGEELRAHPVFAALDTVYALDRHLDRPLREIV 1627
QY 102 GAA-----GLVP-----GMRALKH-GRITLANK 125
DB 1528 AAGEELDTAVTOPALFAFEVALFRLEHGLVDPDLTGHSVGEIAAHVAGVSLDDAA 1687
QY 126 SLVTAGOLMRTAOENATILPVDSHSAVFOALAGEDTACVERVITTAGSGP----- 179
DB 1688 RLVTARGLMOSAREGA-MIAVAGAEVAVESLKGE-----GRVAAVAVGPTAVVYSG 1742
QY 180 -RDWSLERIRACTVAEAOAHNMS-MGQRI-----SIDASMFNKALELIERREFEGF 230
DB 1743 DAD-AAEIRAV-----MAGRGRTRLRVSHAFSPHMDVDLDEFRLVAEGITF 1791
QY 231 EPRDIEAVHPQSTVHAMV-----GFCDDGLMAHLGP 262
DB 1792 EEPRIPIV---STVTGALVTSGETLSPAYWVDQIRRPVRFDAVRTLAADATVLEIGP 1848
QY 263 ADMRAHAGFALNMPGRGEVPAVIDLAQIASLTQKPDDE----- 302
DB 1849 DAVITLAAEALAGT-DAPDAR-DVTIVPRLLRAGREPELTAAGLATAHVHGAPLDRAS 1906
QY 303 -----RPPALRLANDVMAARG--LSGA 322
DB 1907 FFPDGRRTDPTTAFRRHHWLTPEARTDARALGFDPAHRPLLTITVEVAGDGVLLTGR 1966
QY 323 AFNAKEIALDHFITAGRI-----GFLDMAVVEETLAGVSTDPLEFGVPDALEE 371
DB 1967 LSLTDQWMLADHMVNGAVLLPATATFELALAAAGDHVAGVAVVEELTLEAPLVLP 2020

```

```

RESULT 13
US-10-014-717-5
; Sequence 5, Application US/10014717
; Publication No. US20020192778A1
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goelach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/10/014,717
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US/09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0

```


SEQ ID NO 5
LENGTH: 7257
TYPE: PRP
ORGANISM: Sorangium cellulosum
US-10-014-717-5

Query Match 5.7%; Score 113.5; DB 9; Length 7257;
Best Local Similarity 24.0%; Pred. No. 2.4;
Matches 95; Conservative 42; Mismatches 136; Indels 123; Gaps 19;

OY 4 LIFGATGSGESTEDLVNKKGPPEAFRTVALTGGRNIRRLAEMARALAEIA-----VT 58
DB 5510 VESFVSG-----TNAHVLEEAAPAAFAAPAAARSALFVLSASAALDAQARLSAHV 5565
OY 59 AHEDC-LPRLREALGCTEV-----AGGAQAIAEADRPADWTMSAIVGAAGLVPGM 110
DB 5566 AHPELGLGLAFSLATRTSPMYRLAVALTSREALSALD-----TAAQGAAPPA 5615
OY 111 RALKHRTLALANKESLVAGOLIMRTAENGATILPVDESHSAVFOALAGED----- 163
DB 5616 AARGHAST-GSAKVFVFPQO-----GSQWIGMOKLSEEPVPRDAL 5658
OY 164 TACVERVITTAGSGPRDMSLERIRACTVAEQAHPNWSMGORISIDSASMENKALELIE 223
DB 5659 SAC-DEAIOAEAG-----MSL-----LAELADETTSQIGRIDVYOPALFAIEVALSA 5705
OY 224 TREFEFGEDRIEAVVHPQSIYHAMVFGDDGGLMAHLGADMRHAIAGFALNMPGREGVEV 283
DB 5706 LMRWGWEPDAV-----VGISM-----GEVAA 5727
OY 284 ARIDLAQISLTFQKDEREPFALRLADVMARGLSGAFAFNAKETALDHFITAGRIGFL 343
DB 5728 AHV--AGALSL-----EDVAI--ICRSILLRISGGEAMVAVLSLAENALALLGYE 5777
OY 344 DMAAVEETLAGYSTDP-----LEGVADALEVYLA 375
DB 5778 DRLSV-----AVSNSPRSTVLAGE--PALAEVLA 5806

RESULT 14

US-09-308-207-67
Sequence 67, Application US/09308207
Publication No. US20030022323A1
GENERAL INFORMATION:

APPLICANT: MARIA DIAZ-TORRES ET AL.
TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
PRODUCTION OF 1,3 PROPANEDIOL

NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genencor International, Inc.
STREET: 4 Cambridge Place
CITY: Rochester
STATE: NY
COUNTRY: U.S.A
ZIP: 14618

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/308,207
FILING DATE: 13-May-1999
CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/030,601
FILING DATE: 13-NOV-1996
ATTORNEY/AGENT INFORMATION:

NAME: Glaister, Debra
REGISTRATION NUMBER: 33,888
REFERENCE/DOCKET NUMBER: GC 369-2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-864-7620
TELEFAX: 650-845-6504
TELEX: <unknown>
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 609 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: No. US20030022323A1e
SEQUENCE DESCRIPTION: SEQ ID NO: 67:
US-09-308-207-67

Query Match 5.4%; Score 106.5; DB 9; Length 609;
Best Local Similarity 23.5%; Pred. No. 0.32;
Matches 96; Conservative 48; Mismatches 116; Indels 149; Gaps 22;

OY 27 PEAFRTVALTGGRNIRRLAEMARAL-----KAEIAYVAHED-----CLPALREALAGTG--- 75
DB 214 PGIATVFDLMEESQNIYPMARALIGNSAVVYKTPSGDYKRAIPAGLILLIAGRSV 273
OY 76 -TEVAGGAQAIAPADRPADWTMSAIVGAAGL-VPGMRALKHRTLALANKESLVTAGOL 133
DB 274 QYDVAAAGAEAIKAVD--GCGKLDNVAGEAGTNGIM--LEHVRO-----TMAEL 319
OY 134 LMRTAENGATILPVDESHSAVFOALAGEDTACVERVITTAGSGPRDMSLERIRACTVA 193
DB 320 TNKPAQE-----IRIODLLAVDTA---VPVSVTGLAGEFSLEO----- 355
OY 194 EAQAHPNWSMGORISIDSASMENKALELJETREFEFGEDRIEAVVHPQSIYHAMVFGCD 253
DB 356 -----AVGISM-----VKSQRQMALIAETIEKL----- 381
OY 254 GGLMAHLGADMRHAIAGFALNMPGREGVEPYARIDL-----AQIASLTFQKDEREPFAL 307
DB 382 -QIAVOVGGAEBARAILGALTTPGTTR-PLAIDLAGSTDAIIN----- 425
OY 308 RLADYMAARGISGA-----AFNAKETAL-DHFITAGRIGFLDMAAVEETL----- 353
DB 426 --AAGETISATHLAGADDMYMIITARELGLEDRIAEIKKPYLAKV--ESLFIHLRHEDGS 481
OY 354 -----AGVSTD---PLEGVADALEVYLAHMLIARRAE 384
DB 482 VQFPSPALPPTVFARVCVKPDELVPGLDP--LEKVRAL-----RRSAK 524

RESULT 15

US-09-712-363-262
Sequence 262, Application US/09712363
Patent No. US20020164588A1
GENERAL INFORMATION:

APPLICANT: Eisenberg, David
APPLICANT: Rotstein, Sergio H.
APPLICANT: Marcotte, Edward M.

TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
FILE REFERENCE: 07419-032001

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/712,363
FILING DATE: 2000-11-13

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US00/02246
FILING DATE: 2000-01-28

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/179,531
FILING DATE: 2000-02-01

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/117,844
FILING DATE: 1999-01-29

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/118,206
FILING DATE: 1999-02-01

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/126,593
FILING DATE: 1999-03-26

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/134,093
FILING DATE: 1999-05-14

PRIOR APPLICATION NUMBER: 60/165,124
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 60/165,086
PRIOR FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 292
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 262
LENGTH: 1616
TYPE: PRT
ORGANISM: Mycobacterium tuberculosis
US-09-712-363-262

Query Match: 5.48; Score 106.5; DB 9; Length 1616;
Best Local Similarity 26.68; Pred. No. 1.2;
Matches 91; Conservative 38; Mismatches 116; Indels 97; Gaps 20;

QY 89 ADPPADWTMSAIVGAGIVPGMRALKH-----GRTLALANKESIVTACQLMRTAOENG 142
DB 1222 ADRLADGTV-VITGATGAVGVLA-RHLVGAIVGRHVLASR-----RGDRREG 1268
QY 143 ATILPVD-SEHSVAVQALAGE--DTACVE-----RVITASGGGPRD--WS 183
DB 1269 AAELADLTGAKQVYACVADRAAVAGLFAQLSREYPPVGVINHA-GVLDDAVITS 1327
QY 184 LERIRACTVAENQAHPRNMSMQRISIDSASMFNKALELIERFFGPEPPRIEAVHP-Q 242
DB 1328 LTPDRIDTVLAKVDAAWMNLQATSDLDSMFALCSSIAT-----VGSFGQ 1374
QY 243 SIVHAWGFCDDGLMAHLGPDAMRHAIGFALNW-----PGRGEVVARIDLAQIASLTFQ 297
DB 1375 GNYSAANAFLD-GLAAHQAAAGL---AGISLAWGLMEQPGMTAHLSTRDLA----- 1422
QY 298 KPDEERPPALRLANDVNAARGLSGAA--FNAAKEIALDHLFTAGRIIGFLDMAV----- 348
DB 1423 -----RMSRGLAPMSPAEVAFEDAA--LAIDHPLAVAT-LDDRALDARAQAG 1469
QY 349 -VEETLAGVSTDPLEFGVPPDALEVLAMDHLARR-----AAEE 385
DB 1470 ALPALFSGLARPRRRQIDTGDATSSKSALAQRHLGLAADE 1511

Search completed: April 15, 2003, 14:03:58
Job time: 28 secs

[illegible]

```

; FEATURE:
; OTHER INFORMATION: ORF2
US-09-934-903-3
Query Match      19.0%: Score 224.2; DB 10; Length 1182;
Best Local Similarity 56.5%: Pred. No. 2.5e-43;
Matches 446; Conservative 0; Mismatches 328; Indels 15; Gaps 1;

Db      403 ATCTTCATGACGAGCGCTGACGATTCGCGCCCTGTGTGCTGCGATAGACGAGGAC 462
QY      457 TCCGGGCTTTCAAGCGCCGCGGCGAGGACAGCGGCTCG-----CTC 501
Db      463 AAGCGCATCTTTCAGTGCATCCCGCGGCTTATACGCCAGCCATACAGCAACAGCG 522
QY      502 GAGCGGCTATCATACAGCGCTCCGCGGCGGCTTCCGCGATGAGCCCTGAGCGCATC 561
Db      523 CGCGGCATTTATTTAGCCGCTTCCGCGTCCGCTTTGAGAGGAGCGCCATAGAAACG 582
QY      562 CGCGCTGACACGCTGCGGAGCGGAGCCCATCCCAATGCTTCATGCGGCGAGCATC 621
Db      583 TCCAGCGCTACAGCGGATACAGCGCTTCCGCTTAAATGAGCATGCGGCGCAAGATT 642
QY      622 TCCATGACAGCGGCTGATGTTCAACAGGCGCTGCGATGATGAGAGCGGCAATTC 681
Db      643 TCGGTCGATTTCCGCGCATCATATGACAAAGGCTTCGACTGATGAGAGCTGCTTGT 702
QY      682 TTGCGCTTCAGCGCGGAGCGGATGAGCGGCTGCTCATCCGCAATCCATGCTCATG 741
Db      703 TTCAACATGAGCGCGGACCATGATTAAGTGTGATTCGCGAGAGCATCATTCATTCG 762
QY      742 ATGCTGGCTTCTGCGAGGCGGCGCTGATGCGCCATCTCGGCGCGCGCATGCGCAC 801
Db      763 ATGCTGACTATGCTGATGCTGCTGCTTTGGCGAGATGGGTAAATCCGACATGGCG 822
QY      802 GCCATCGGATTCGCGCTGACCTGCGCGGCGGCGGCGGCGCGCGGCGGATTCAC 861
Db      823 CCGATAGCGGACGCGGATGCGCTGCGCGGAGACGCTTGTGCTGCTGCTGCTGAT 882
QY      862 CTGCGACAGATTCGAGCGCTACCTTCAGAGCGCTGACAGAGAGCGTTTCCGCGCTG 921
Db      883 ATTTTCGAGTAGGCGCATATGATTTTCGAAACCCGACTTGAAACGGTTTCTTGTCTG 942
QY      922 AGGCTTGGCGGAGAGCGATGATGGCGGCGGCGGCGGCGGCGGCGGCTTAAAGCGG 981
Db      943 AGATTGGCTTATGAGCGCATATAGCTGTGTGAATATATGCCAACGGATTAACGCGAC 1002
QY      982 AAGAGATCGCGCTGATCATTTTCATGCGCGCGAGCATGCGGCTTCTGAGCATGCGCG 1041
Db      1003 AATGAATTCGCTGCGAAGCGTTTAAATGAGAGATCAATTCACATGACATGCGGCTC 1062
QY      1042 GTGCTCGAG 1050
Db      1063 ATCATCGAG 1071

```

```

; FEATURE:
; OTHER INFORMATION: ORF2
US-09-934-903-3
Query Match      19.0%: Score 224.2; DB 10; Length 1182;
Best Local Similarity 56.5%: Pred. No. 2.5e-43;
Matches 446; Conservative 0; Mismatches 328; Indels 15; Gaps 1;

QY      277 GCCGCTGACGATGTCGCGCATGTCGCGGCGCGCGGCGCTGCGCGGCAATGCGGCG 336
Db      283 GTGCTGATGCGATGCGCGGCTATGCTGCGCGGCGCGGCGGCTGCTGCGGCGGCG 342
QY      337 CTGAAGCAGCGCGCGACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 396
Db      343 GCCAAGCGCGCGCAAAACCGCTGCTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 402
QY      397 CTCTGATGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 456
Db      403 ATCTTCATGACGAGCGGCTGACGATTCGCGGCTGCTGCTGCTGCGGATAGACAGG 462
QY      457 TCCGCGGCTTTCAGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 501
Db      463 AAGCGCATCTTTCAGTGCATCCCGCGGCTTATACGCCAGCCATACAGCAACAGCG 522
QY      502 GAGCGGCTATCATACAGCGCTCCGCGGCGGCGGCTTCCGCGCATGAGCGCTGAGCGCATC 561
Db      523 CGCGGCATTTATTTAGCCGCTTCCGCGTGGCGCGGCGGCGGCGGCGGCGGCGG 582
QY      562 CGCGGCTGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 621
Db      583 TCCAGCTGACGCGGCGGATCAGGCGGCTGCGCATCTTAATGAGCATGAGGCGCAAGATT 642
QY      622 TCCATGACAGCGGCTGATGTTCAACAGGCGCTGCGAGCTGATGAGAGCGGCAATTC 681
Db      643 TCGGTCGATTCGCGCATCATATGATGACAAAGGCTTCGAACTGATGAGAGCGCTG 702
QY      682 TTGCGCTTCAGCGCGGAGCGGATGAGCGGCTGCTCATCCGCAATCATGCTCATGCG 741
Db      703 TTCAACATGAGGCGCGGACCATGATTAAGTGTGATTCGCGAGAGCATATTCATTCG 762
QY      742 ATGCTGGCTTCTGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 801
Db      763 ATGCTGACTATGCTGATGCTGCTGCTTTGGCGAGATGGGTAAATCCGACATGCGCAG 822
QY      802 GCCATCGGATTCGCGCTGACCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGATTCAC 861
Db      823 CCGATAGCGGACGCGGATGCGCTGCGCGGAGACGCTTGTGCTGCTGCTGCTGAT 882
QY      862 CTGCGACAGATTCGAGCGCTACCTTCAGAGCGCTGACAGAGAGCGTTTCCGCGCTG 921
Db      883 ATTTTCGAGTAGGCGCATATGATTTTCGAAACCCGACTTGAAACGGTTTCTTGTCTG 942
QY      922 AGGCTTGGCGGAGAGCGATGATGGCGGCGGCGGCGGCGGCGGCGGCTTAAAGCGG 981
Db      943 AGATTGGCTTATGAGCGCATATAGCTGTGTGAATATATGCCAACGGATTAACGCGAC 1002
QY      982 AAGAGATCGCGCTGATCATTTTCATGCGCGCGAGCATGCGGCTTCTGAGCATGCGCG 1041
Db      1003 AATGAATTCGCTGCGAAGCGTTTAAATGAGAGATCAATTCACATGACATGCGGCTC 1062
QY      1042 GTGCTCGAG 1050
Db      1063 ATCATCGAG 1071

```

```

RESULT 2
US-09-934-903-3
; Sequence 3, Application US/09934903
; Patent No. US20020102690A1
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mathews
; APPLICANT: Odom, J. Martin
; APPLICANT: Schenzle, Andreas J.
; APPLICANT: No. US20020102690A1ton, Kelley C.
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Rouviere, Pierre
; APPLICANT: Plicataggio, Stephen
; APPLICANT: Cheng, Olong
; TITLE OF INVENTION: Genes Involved in Isoprenoid Compounds Production
; FILE REFERENCE: C11646 US NA
; CURRENT APPLICATION NUMBER: US/09/934,903
; PRIOR APPLICATION NUMBER: 2001-08-22
; PRIOR FILING DATE: September 1, 2001
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 1182
; TYPE: DNA
; ORGANISM: Methylobionas 16a

```

```

RESULT 3
US-09-934-868-63
; Sequence 63, Application US/09934868
; Patent No. US20020137190A1
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mathews
; APPLICANT: Odom, James M
; APPLICANT: Schenzle, Andreas J

```

```

: TITLE OR INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
: FILE REFERENCE: C11596 US NA
:
: CURRENT APPLICATION NUMBER: US/09/934,868
: CURRENT FILING DATE: 2001-08-22
: PRIOR APPLICATION NUMBER: 60/229,858
: PRIOR FILING DATE: 2000-09-01
: NUMBER OF SEQ ID NOS: 81
: SOFTWARE: Microsoft Office 97
: SEQ ID NO 63
:
: LENGTH: 1182
: TYPE: DNA
: ORGANISM: Methylomonas 16a
: FEATURE:
:
: OTHER INFORMATION: DXR
:
: US-09-934-868-63

```

Query Match	19.0%;	Score 224.2;	DB 10;	length 1182;
Best Local Similarity	56.5%;	Pred. No. 2.5e-43;		

[illegible]

Db 1063 ATCATCGAG 1071

```

RESULT 4
US-09-938-956-5
; Sequence 5, Application US/09938956
; Patent No. US20020142408A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Siqun
; APPLICANT: Dicosimo, Deana J.
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odcom, J. Martin
; TITLE OF INVENTION: Production of Monoterpene
; FILE REFERENCE: CIL809 US NA
; CURRENT APPLICATION NUMBER: US/09/938,956
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/229,907
; PRIOR FILING DATE: 2000-09-0
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 11575
; TYPE: DNA
; ORGANISM: Plasmid
; US-09-938-956-5

```

Query Match	19.08:	Score 224.2:	DB 10:	Length 11575:
Best Local Similarity	56.58:	Pred. No. 2e-43:		
Matches 446:	Conservative	0:	Mismatches 328:	Indels 15:
				Gaps 1

QY	277	GCACACGAGCAACATGTCGGCCATGTCGTGGGGCGCGGGGGCTCGTCGGGATAAGGGGAGG	336
Db	3339	GTCCTATACGGGTGATGGGGCTATGCTGTGGGGCGGGCGGATGTGGTCCGACTTGGCCGGG	3399
QY	337	CTGAGACACGCGCGCACGCGTGGGGCTCGCCACAAAGAAAGCCTCGTGCAGCGCAGAGCA	396
Db	3399	GCCAAAGCGCCGCAAAACCGTGTGCTGGTGGCCAAACAAAGAAAGCCTGTGTATGTGCGGAGCA	3459
QY	397	CTCCTGATGCGGAGAGCGCCACAGGAAGAGGGCGCCAGATCCTGTCCGTGGACAGCAGCAC	456
Db	3459	ATCTTCATGACAGCGCCGTCACACGATCTCCGGCGCTGTGTCTCGCCGATACAGCAGCAGCAC	3518
QY	457	TCGCCGGTCTTTCAAGGGCTGGGGGGCGAGAACAGCGCTGC-----GTC	501
Db	3519	AACGCCATCTTTTATAGTCATGCGCGGGGGTATATAGCCACAGGCCATATACGCCAAAGAGGG	3578
QY	502	GAGCGGCTCATCATACAGCGGCTCGCGCGGGCCGTTCCGAGACTGAGAGCTCGACGGCATC	561
Db	3579	CGCCGCGATTTTATTGACCGCTTCCGGTGGCCATTTCGACGAGACGCCGATTAAGAACGTTG	3638
QY	562	CGCGCCATGCACCGGGGCGAGGGCGAGGCCCATCCCACTGCTCATAGGGCCAGCGGATC	621
Db	3639	TCCAGCGTCACGCGCGGATCAGGCGCTGTGGCCATCTAAATGGAGCATGAGGGCGCAAGATT	3698
QY	622	TCCATTCACAGCGCGCTCGATGATTCAACAAAGCGCTGAGCTGATGAGAGCGCGCATTC	681
Db	3699	TCGGTCATTTCCSCACACATGATGAAGAACAAAGGCTCTGAACTATGCAACCCCTGCTTGTG	3758
QY	682	TTGCGCTTCGAGCGCGGACCGGATCGAGGGGGGTGTCATCCGCAATCCATCTGTCATGAG	741
Db	3759	TTCAACATGAGAGCCCGACGACGATTTGAATGTGCTATTCAATCCGACAGCATCTTATTCG	3818
QY	742	ATGTGAGGCTTTCGCAAGGGGGCGTGATGGCCCATCTCGGCCCGCGCGCATATGCGCAC	801
Db	3819	ATGTGTGACATATGTCATGTGGTTGGATTGGGCGCAATGGGTAATCCGACATGCGGCACG	3878
QY	802	GCCATCGGATTTCCCGCGTGAACATCGCGGGGTCCGGCGAGAGTCCCGTGCCCGGATTCAC	861
Db	3879	CCGATTACGACGACGAGTGGCGCTGGCCGGAACGCTTTGACTCTGTGTGGCGCCGCGCTGAT	3938

QY 862 CTCGCACGATTCGAGCGCTTACCTTCAGAGCCTGACAGGAAGCTTTCCGGCCCTG 921
 Db 3939 ATTTTCGAGTAGGGGACATGATGATTTGAAAAACCCGACTTGAACGGTTCTTGTCTG 3998
 QY 922 AGCGTTGGCGAGACGTCATGATGCGCGCGCCGCTGTGGGGCCGCCCTTCAACGGCCG 981
 Db 3999 AGATTGGCTTATGAAAGCATCAAGTCGTGTGATGATTAATGCAACGGTATTGAACGACCC 4058
 QY 982 AAGGAGATCGCGCTGATCATTTTCATCCGCGAGCAGCATGGGTTTCTGACATGGCGCG 1041
 Db 4059 AATGAATTTGTCGTGAGCGCTTTTAAATGAAGAATCAATTCACATCGCATCGCGTTC 4118
 QY 1042 GTGGTCGAG 1050
 Db 4119 ATCATCGAG 4127

RESULT 5

US-09-712-363-112
 : Sequence 112, Application US/09712363
 : Patent No. US20020164588A1
 : GENERAL INFORMATION:
 : APPLICANT: Eisenberg, David
 : APPLICANT: Rotstein, Sergio H.
 : APPLICANT: Marcotte, Edward M.
 : TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
 : FILE REFERENCE: 07419-032001
 : CURRENT FILING DATE: 2000-11-13
 : PRIOR APPLICATION NUMBER: PCT/US00/02246
 : PRIOR FILING DATE: 2000-01-28
 : PRIOR APPLICATION NUMBER: 60/179,531
 : PRIOR FILING DATE: 2000-02-01
 : PRIOR APPLICATION NUMBER: 60/117,844
 : PRIOR FILING DATE: 1999-01-29
 : PRIOR APPLICATION NUMBER: 60/118,206
 : PRIOR FILING DATE: 1999-02-01
 : PRIOR APPLICATION NUMBER: 60/126,593
 : PRIOR FILING DATE: 1999-03-26
 : PRIOR APPLICATION NUMBER: 60/134,093
 : PRIOR FILING DATE: 1999-05-14
 : PRIOR APPLICATION NUMBER: 60/134,092
 : PRIOR FILING DATE: 1999-05-14
 : PRIOR APPLICATION NUMBER: 60/165,124
 : PRIOR FILING DATE: 1999-11-12
 : PRIOR APPLICATION NUMBER: 60/165,086
 : PRIOR FILING DATE: 1999-11-12
 : NUMBER OF SEQ ID NOS: 292
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 112
 : LENGTH: 1311
 : TYPE: DNA
 : ORGANISM: Mycobacterium tuberculosis
 : US-09-712-363-112

Query Match 18.2%; Score 215.6; DB 9; Length 1311;
 Best Local Similarity 54.2%; Pred. No. 2.4e-41;
 Matches 462; Conservative 0; Mismatches 384; Indels 6; Gaps 1;

QY 201 GCGGAGGCGGTGGCGGAGCGGACCGAGGTGCGCGGGGGGCGGACGAGCGCATGGCCGA 260
 Db 282 GCACGCGGCGGAGGCGGTGGCGGACATCCCTACCGAGATCCGACGCGCCACCCGGCT 341
 QY 261 GCGCGCGACCGCGCGCGGACGATGCGGACATGCGGACATGCGGCGCGCGGCGCTGT 320
 Db 342 GTGTGAGAGACGACGCGGACGTCGTCTCAATGCGTGTGTCGCGCGGTGGGCTCGG 401
 QY 321 GCGCGGAATGCGGCGGTGAACGACGCGCGGACGCTGGCGCTGCGCAACGAAGAAAGCT 380
 Db 402 ACCGACGTTGGCGCGGTCACAGACGCGGTGCGCGCTGCGCTGCGCAACGAAGATCGCT 461
 QY 381 CGTAGCGGACGAGCAATCTCTGATGCGGACGCGCCAGAGAAAGCGCGCACGATCTCTCC 440

Db 462 GGTCCCGGTGTTGCTGT 515
 QY 441 GGTGAGACGCGACGACCTCCGCGGCTTTTCAAGCCGTGCGGCGGACGACGACGCGCTGCT 500
 Db 516 GGTGACATTCGACACCTCCGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 575
 QY 501 CGAGCGCTCATCATCAGCGGCGTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 560
 Db 576 CGCCAGCTGT 635
 QY 561 CCGCGCTGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 620
 Db 636 CGAGCATGTCAACCCCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 695
 QY 621 CTCATCGACAGCGCGCTGATGTTCAACAGGCGCTGAGCTGATCGAGACGCGGCAATT 680
 Db 696 CACGCTGAATTCGCGGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 755
 QY 681 CTTGCGCTTTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 740
 Db 756 GTTCGCGATCCCTTCACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 815
 QY 741 GATGTGGGCTTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 800
 Db 816 GATGTGACCTTCATCATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 875
 QY 801 CGCCATCGGATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 860
 Db 876 ACCGATTTCTGTATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 935
 QY 861 CCTGCGACAGATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 920
 Db 936 TTTCCATACCCGCTGTCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 995
 QY 921 GAGCGTTCGCGGAGAGCGTCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 980
 Db 996 CGAGTTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1055
 QY 981 CAAGAGATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1040
 Db 1056 GACGAGAGACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1115
 QY 1041 GGTGTGCGAGA 1052
 Db 1116 CATATCGCGGA 1127

RESULT 6

US-09-974-300-1083
 : Sequence 1083, Application US/09974300
 : Patent No. US20020146721A1
 : GENERAL INFORMATION:
 : APPLICANT: Berka, Randy M.
 : APPLICANT: Clausen, Ib Groth
 : TITLE OF INVENTION: Methods For Monitoring Multiple Gene
 : FILE REFERENCE: 10085.500-US
 : CURRENT APPLICATION NUMBER: US/09/974,300
 : PRIOR FILING DATE: 2001-10-05
 : PRIOR APPLICATION NUMBER: 09/680,598
 : PRIOR FILING DATE: 2000-10-06
 : PRIOR APPLICATION NUMBER: 60/279,526
 : PRIOR FILING DATE: 2001-03-27
 : NUMBER OF SEQ ID NOS: 8481
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 1083
 : LENGTH: 1115
 : TYPE: DNA
 : ORGANISM: Bacillus licheniformis
 : US-09-974-300-1083

Query Match 15.2%; Score 179.2; DB 10; Length 1115;

Best Local Similarity 50.0%; Pred. No. 6,4e-33;
Matches 535; Conservative 0; Mismatches 523; Indels 12; Gaps 3;

```
QY 2 TGGGAGCCTGTCGATCTTTGGGGCCACCGGCTCCATCGCGCAATCCACTTCGACCTCG 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2 TGAAGAACTATGTCATTTGGAGGACAGGCTCCATAGAGCAACAGACGCTTGATGATCA 61
QY 62 TCATCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 62 TCAAAAGCGATG-----ACGATAGTTTCGGCTGACCCGCTGACTTTCGGCAAAAACG 115
QY 122 TCGGCGACTGCGCGAATGCGCGCTGAGGCGAGGCTTGGCGTACCGCGCATG 181
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 116 CCGAAAAACGGGAGAAATCATTAACCTTTAAACCGAAGATATGAGCGTCGGGAGATG 175
QY 182 AGAGTGCCTGCGCGCGCTGCGGAGGCGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 241
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 176 AGCATATCATATGAAACATTAAACAGCATTTCTTCCTACACCTTTAAACAGGATCG 235
QY 242 GGGCGAGGCG---CATCGCGAGGCGCGGCGCGGCGCGGCGCGATGTCGGGCA 298
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 236 GGGAGAGAGCTTTAATCGAAGCGGCGCTTATCCGGAAGCGATATCGTCATATGCTC 295
QY 299 TCGTGGGCGCGCGGCGCTCGTCCCGGAATGCGGCGCTGAAAGCAGGCGCGCGCTG 358
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 296 TCGTGGGCGCGCTGCTTTCTCCGACTTTAAAGCGGATGAGCATTAATAAACAAATG 355
QY 359 CGCTGGCCAAAGAAAGAGGCTCGTGCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 418
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 356 CGCTTGCATTAATGAAACATCTGTCGACGCGGCTCATTAAGAAAGAAACAGCCCAA 415
QY 419 AGAAGCGGCCAGATCTCGCGGTGACAGGAGCATCTCGCGGCTTTCAGAGGCGCTG 478
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 416 AATAGATATCTCCGCTCGCGCTCGCGCTGACAGGAGCAATTCGCGATTTCCAGGCGCTTC 475
QY 479 CGGCGAGGAGACAGGCGCTGCTGAGCGGCTCATCATCAGGCGCTGCGGCGCGCTTC 538
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 476 AGGCGGAAATATCGAAAGCATATCAAAAGGCGATGCTGACGCGTTCAGCGGAGCTTCA 535
QY 539 GGGAGGAGGCTCGAGGCGCATCCGCGCTGACAGGCGGAGGAGGAGGAGGAGGAGGAGG 598
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 536 GGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 595
QY 599 ACTGTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 658
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 596 ATTGTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 655
QY 659 AGCTGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 718
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 656 AGGTTATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 715
QY 719 ATCCGCAATCCATGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 778
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 716 ACAAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 775
QY 779 TCGGCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 838
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 776 TTGGGACACCTGATATGAGAGGCGGATCATATGCGCTGACGCAATCCGGAAGCGCTCG 835
QY 839 AGGTGCGCGTGC---CGGATGAGCTTGCACAGATTGCGAGCTCACCTTCGCAAGC 895
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 836 CATTAAATGAGCAAAATCCCTGATCTCTGGGAAGTCCGAGACGCTGATTTTGGCCAAG 895
QY 896 CTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 955
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 896 CTGATTTTGAAGAGTTCCCTGCTGTACAAATGCTTATATAGTAATAGGAGGAGTA 955
QY 956 TGTGGGCGCGCGCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1015
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 956 CAATGCGGAGCTGTTTAAACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1015
QY 1016 GCATCGGCTTTTGACATGCGGCGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1065
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

DB 1016 GGATTTCTTCTTGGAATAGAGACATCATTTGAAAAGGCTTTGGAACGC 1065

RESULT 7

US-09-738-626-2209

; Sequence 2209, Application US/09738626

; Publication No. US20020197605A1

; GENERAL INFORMATION:

; APPLICANT: NAKAGAWA, SATOSHI

; APPLICANT: MIZOGUCHI, HIROSHI

; APPLICANT: ANDO, SEIKO

; APPLICANT: HAYASHI, MIKIRO

; APPLICANT: OCHIAI, KEIKO

; APPLICANT: YOKOI, HARUHIKO

; APPLICANT: TATEISHI, NAOKO

; APPLICANT: SENOH, AKIHIRO

; APPLICANT: IKEDA, MASATO

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-125

; CURRENT APPLICATION NUMBER: US/09/738,626

; CURRENT FILING DATE: 2000-12-18

; PRIOR APPLICATION NUMBER: JP 99/37484

; PRIOR FILING DATE: 1999-12-16

; PRIOR APPLICATION NUMBER: JP 00/159162

; PRIOR FILING DATE: 2000-04-07

; PRIOR APPLICATION NUMBER: JP 00/280988

; NUMBER OF SEQ ID NOS: 7059

; SOFTWARE: PatentIn ver. 3.0

; SEQ ID NO 2209

; LENGTH: 1176

; TYPE: DNA

; ORGANISM: Corynebacterium glutamicum

US-09-738-626-2209

Query Match

Best Local Similarity 12.3%; Score 145.8; DB 9; Length 1176;

Matches 397; Conservative 0; Mismatches 392; Indels 6; Gaps 1;

```
QY 276 GGGCGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 335
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 285 GGGCGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 344
QY 336 GCGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 395
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 345 TCTGATATGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 404
QY 396 ACTCTGATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 455
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 405 GTTGTATCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 458
QY 456 CTCGCGGCTTTCAGGCGCTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 515
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 459 CTCGCGGATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 518
QY 516 CAGGCGGCTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 575
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 519 GACAGCTTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 578
QY 576 GGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 635
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 579 CGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 638
QY 636 CTCGATGTTCAACAGGCGCTGAGCTGATCGAGAGCGCGGAGTTCCTTGGCTTGAGCGC 695
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 639 CACCTTATTAATAAGGCTCGAATCTCATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 698
QY 696 GAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 755
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 699 GATCTCATTTGATGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 758
QY 756 CGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 815
```

Query Match	12.3%	Score 145.8	DB 9:	Length 3309400
Best Local Similarity	49.9%	Pred. 1.7e-25		
Matches 397	Conservative	0	Mismatches 392	Indels 6
			Gaps 1	
QY 276	GGCCGACGTGACCAATGTCGCCATCGTGGGCGCGCGGCTCTGTCGCCGGAATCGGGC	335		
Db 2130794	GGCCGACACATGTGCTTAATCTCTGTTGGTCTTTTGGGCTTCCGGACACCGTGGCCAC	2130735		
QY 336	GCTGACGACGCGCGCGACCGCTGGCGGCTGCCAACAAGAAAGCTCTGACGGCAGGGCA	395		
Db 2130734	TCTGGAAATCTGTCGCGCATCTTGCTTGCGCTAACAAAGAAATCGCTGTTGCGGCTGTGA	2130675		
QY 396	ACTCCTGATGGGAGCGCGCCACAGAGAAAGGGCGCACATCTGCCCGGTGGACAGCAGACA	455		
Db 2130674	GTTTGTACTCTCAAAAGCCAAAGCTG-----GGCGCAATCATTTCCGTCATTCGGAGCA	2130621		

```

RESULT 9
US-09-923-556-5
; Sequence 5, Application US/09923556
; Patent No. US20020058321A1
; GENERAL INFORMATION:
; APPLICANT: Busch, Marco
; APPLICANT: Hahn, Rudiger
; APPLICANT: Martin, William
; APPLICANT: Tietjen, Klaus
; APPLICANT: Klotz, Andreas
; TITLE OF INVENTION: Method of determining the activity of
; TITLE OF INVENTION: 1-deoxy-D-xylylucose-5-phosphate synthase
; TITLE OF INVENTION: 1-deoxy-D-xylylucose-5-phosphate synthase
; FILE REFERENCE: 2020US
; CURRENT APPLICATION NUMBER: US/09/923,556
; CURRENT FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/449,335
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1732
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (18)..(1448)
; OTHER INFORMATION: new coding sequence

```



```

; APPLICANT: Busch, Marco
; APPLICANT: Hain, Rudiger
; APPLICANT: Martin, William
; APPLICANT: Tietjen, Klaus
; APPLICANT: Rieth, Andreas
; TITLE OF INVENTION: Method of determining the activity of
; TITLE OF INVENTION: 1-deoxy-D-xylulose-5-phosphate reductoisomerase and
; TITLE OF INVENTION: 1-deoxy-D-xylulose-5-phosphate synthase
; FILE REFERENCE: 202005
; CURRENT APPLICATION NUMBER: US/09/923,556
; CURRENT FILING DATE: 2001-08-07
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: US/09/449,335
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1732
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (18)..(1448)
; OTHER INFORMATION: DNA encoding 1-deoxy-D-xylulose-5-phosphate
; OTHER INFORMATION: reductoisomerase
US-09-923-556-1

```

```

Query Match      10.5%; Score 123.6; DB 10; Length 1732;
Best Local Similarity 48.7%; Pred. No. 4.5e-20;
Matches 373; Conservative 0; Mismatches 384; Indels 9; Gaps 1;

```

```

QY 293 CGGCCATCTGGGCGCGCGGCGCTGCTGCGGAGTGCAGGCGCTGAGCAGCGCGCA 352
DB 550 CCGGAATAGTAGTGTGCGGAGCTAAAGCTTACGCTGCTGCAATTAAGCAGGAAG 609
QY 353 CGGTGGCGCTCGCCCAAGAGAAAGCTCTGAGCGGCGCAACCTCTGATGCGGAG 412
DB 610 ACATGCTCTTGCACAAAGAGACATTAATCCAGGTGCTCTTGTGCTTCCGCTTG 669
QY 413 CCCAGGAGAAAGCGCGCAGCATCTCTGCTGAGCAGGAGCAGCTCGCGGCTTTGAG 472
DB 670 CCAACAAACATTAATGTAAATTTCTTCCGCGAGATTCAGAAATTTGCCATATTTCAGT 729
QY 473 CGGTGGGCGGAGGAGACAGGCTGCTGAGCGGCTCATCATCACGCGCTCCGCGGCG 532
DB 730 GATTTCAAGGTTTGCCTGGAAGGCGCTCTGCGCAAGATTAATCTTGACATCTGCTGGAG 789
QY 533 CGTTCCGCGAGTGGAGCGCTGAGCGCATCCGCGCTGACCGTGGCGGAGCGAGCGCC 592
DB 790 CTTTGGAGGATTTGGCCCTGTGAAAGAGCTAAAGAAAGTTAAAGTAGCGGATGCGTTGAGC 849
QY 593 ATCCCAACTGGTCCATGCGGCGAGGATCTCATTCGACAGCGCGCTCGATGTTCAACAGG 652
DB 850 ATCCCAACTGGAGCATGGAAGAAAGAAATCACTGTGACCTCTGACCTTTTCAACAGG 909
QY 653 CGGTGAGCTGATGAGAGCGCGGAATTTCTTGGCTTCGAGCGCGGAGCGGATGAGCGCG 712
DB 910 GTCTTGGAGTATGAGAGCGCATTAATTTGTTGAGCTGAGTATGAGCATATAGAGATTG 969
QY 713 TCGTCCATCCGCAATCATGCTCATGCTGATGCTGAGGCTCTGCGAGCGGCGGCTGATG 772
DB 970 TCAATTCATCGCAAGATATATACATTCATGATGAAACAGAGATTCATCTGTGCTTG 1029
QY 773 CCCATCTCGGCGCGCGGAGCATGCGGCGCATGCGATTCGCGCTGCAACTGCGCGGCTG 832
DB 1030 CTCATTTGGGTTGGCCGATATGCTTTTACCGATTTCTCTACCATGTCATGCGCGGATA 1089
QY 833 GCGGCG-----GAGGTCCCGGTGCGCGCGGAGTGCAGCTCGCAGAGATTGCGAGCTCA 883
DB 1090 GAACTTCCTTGTCTGAAGTACTTGGCCAGAGCTTGGACCTTTCGCAAACTCGGTTCAATTGA 1149
QY 884 CTTTCCAGAGCCTGAGAGAGAGAGCTTTCCGCGCGCTGAGAGCTTGGCGGAGAGCTATG 943
DB 1150 CTTTCAAGAAACGAGCAATGTGAATACCATTCATGATGATCTTGTATGCTGCTGAGC 1209

```

```

QY 944 CGGCGCGGCGCTGCTGCGGCGCGGCTTCAACGCGGCGCAAGAGATCGCTGATCAT 1003
DB 1210 GAGCTGAGGCGCACATGACTGAGATTTCTACAGCGCGCGCATGAGAAAGCTGTGAATGT 1269
QY 1004 TCATTCGCGGAGCATGCGGCTTCTGACATGCGGCGGCTGTGCA 1049
DB 1270 TCATTTGATGAAGAAAGATTAAGCTATTGATGATCTTCAAGGTTTGCA 1315

```

RESULT 12

US-10-047-412A-9

; Sequence 9, Application US/10047412A

; Publication No. US20020197696A1

; GENERAL INFORMATION:

; APPLICANT: Levin, Joshua Z.

; APPLICANT: Potter, Sharon L.

; APPLICANT: Weiglich, Lynette M.

; TITLE OF INVENTION: Herbicide Target Genes and Methods

; FILE REFERENCE: PB/5-30780DIV

; CURRENT APPLICATION NUMBER: US/10/047,412A

; CURRENT FILING DATE: 2002-04-11

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 9

; LENGTH: 1434

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1434)

US-10-047-412A-9

```

Query Match      10.3%; Score 122; DB 9; Length 1434;
Best Local Similarity 48.6%; Pred. No. 1.1e-19;
Matches 372; Conservative 0; Mismatches 385; Indels 9; Gaps 1;

```

```

QY 293 CGGCCATCTGGGCGCGCGGCGCTGCTGCGGAGTGCAGGCGCTGAGCAGCGCGCA 352
DB 550 CCGGAATAGTAGTGTGCGGAGCTAAAGCTTACGCTGCTGCAATTAAGCAGGAAG 592
QY 353 CGGTGGCGCTCGCCCAAGAGAAAGCTCTGAGCGGCGCAACCTCTGATGCGGAG 412
DB 610 ACATGCTCTTGCACAAAGAGACATTAATCCAGGTGCTCTTGTGCTTCCGCTTG 652
QY 413 CCCAGGAGAAAGCGCGCAGCATCTCTGCTGAGCAGGAGCAGCTCGCGGCTTTGAG 472
DB 653 CCAACAAACATTAATGTAAAGATTTCTTCCGCGAGATTCAGAAATTTGCCATATTTCAGT 712
QY 473 CGGTGGGCGGAGGAGACAGGCTGCTGAGCGGCTCATCATCACGCGCTCCGCGGCG 532
DB 713 GATTTCAAGGTTTGCCTGGAAGGCGCTCTGCGCAAGATTAATCTTGACATCTGCTGGAG 772
QY 533 CGTTCCGCGAGTGGAGCGCTGAGCGCATCCGCGCTGACCGTGGCGGAGCGGCGCC 592
DB 773 CTTTGGAGGATTTGGCCCTGTGAAAGAGCTAAAGAAAGTTAAAGTAGCGGATGCGTTGAGC 832
QY 773 ATCCCAACTGGTCCATGCGGCGAGGATCTCATTCGACAGCGCGCTCGATGTTCAACAGG 652
DB 833 ATCCCAACTGGAGCATGGAAGAAAGAAATCACTGTGACCTCTGACCTTTTCAACAGG 892
QY 833 CGGTGAGCTGATGAGAGCGCGGAATTTCTTGGCTTCGAGCGCGGAGCGGATGAGCGCG 712
DB 893 GTCTTGGAGTATGAGAGCGCATTAATTTGTTGAGCTGAGTATGAGCATATAGAGATTG 952
QY 713 TCGTCCATCCGCAATCATGCTCATGCTGATGCTGAGGCTCTGCGAGCGGCGGCTGATG 772
DB 953 TCAATTCATCTCAAAAGATATACATTCATGATGAAACAGAGATTCATCTGTGCTTG 1012
QY 773 CCCATCTCGGCGCGCGGAGCATGCGGCGCATGCGATTCGCGCTGCAACTGCGCGGCTG 832
DB 1013 CTCATTTGGGTTGGCCGATATGCTTTTACCGATTTCTCTACCATGTCATGCGCGGATA 1072

```


Search completed: April 22, 2003, 06:16:53
Job time : 3994 secs

Oy	515	TCACGGCGCTCGGGGGCGCTTCCGACATCGACCTTCAGAGCGATCCGGCCGTCACCG	574
Db	701	GCGCCCTGCGCGAAGAGTGTGTGCCGACGCCCGGGCGGACGCGCGCCGCGCGCGCG	760
Oy	575	TGCGCCAGAGCGACGCCCATCCCACTGCTTCAT	608
Db	761	AGCGCGGCGCGACGCCCGCGGCTGTGCGCGCTT	794

RESULT 15

US-09-860-846-32
Sequence 32, Application US/09860846
Patent No. US20020164742A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding mycymcin and plikromycin
FILE REFERENCE: 600,438US1
CURRENT APPLICATION NUMBER: US/09/860,846
CURRENT FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 32
LENGTH: 11220
TYPE: DNA
ORGANISM: Streptomyces venezuelae
US-09-860-846-32

Query Match	5.68	Score	66.6	DB	9	Length	11220
Best Local Similarity	47.38	Pred.	No. 5.9e-07				
Matches	235	Conservative	0	Mismatches	259	Indels	3
						Gaps	1

QY	156	GGCGAGCTTGCCTGACGCGCGGATGAGAGACTCTGCGCCGGGCTGCGGAGAGCGCTGGC	215
Db	1707	GGTGGCTTGTCTGTTTCCCGGTGAGGGCACGCAAGTGGCCGGAGTGGCGCGCAACTCT	1766
QY	216	CGGAGCGGACCGAGGCTGCGGGGCGGGGCGCAGGCCATTGCCGAGGCGCGACCGGC	275
Db	1767	CGAGGTGTCCAAAGAGATTTCCTCCGGCGGCACATGGCCGAGTGCAGAGGCGCGCTTCCCGTA	1828
QY	276	GGCCGACTGGAGCACTGTGCGCAATCGGGGCGCGGGGCGCTGTCGCCCCGGAATGCGGGC	335
Db	1827	TGTGCACTGGTCCGTGGAGGCGCTGTCGCGCAGGCGCGCGGCGCGCCACGCTGGAGCG	1886
QY	336	GCTGAGACAGCGCGCCGACGCTGGCGCTCGCCAAACAAGAAAGCTTCGTAGCGGACGCA	395
Db	1887	GGTGCAGTGTGCTCAGCCGCTGTGACCTTCGCTCATGTGTTGCTGGCAAGATCTGGCA	1946
QY	396	ACTCTGATCGGGCGGCGCCGAGAGAAAGCGCGCACAGATCTGCGCGTGGACAGGAGACA	455
Db	1947	GCACCAAGCGGTGAGCGCCGCAAGGCGCGCTGCTGGCGCA--CTGCGAGGGCGAATGCGCGC	2003
QY	456	CTCCGCGGTCTTTCAGGCGCGCTGGCGGAGAGACAGGCGCTCGTGCAGCGGCTCATAT	515
Db	2004	CGCTGACGTGCGCGGTGTCCTTCACCTCGACGACGCGCGCGGTGTCACCTCGGCA	2063
QY	516	CAGGGCGTTCGGGGGGCGCTTTCGGGCATGAGACCTTCAGAGCGCATTCGGGCTTGACGCT	575
Db	2064	CAATCCATGCGCGCCCACTCTCCCGGCAAGGGCGGCATGATCTCCCTGCGCTCAGGA	2123
QY	576	GGCGAGGCGAGGCCATCCCAACTGTGTCATGTGGCGCAGCGACTTCATTCAGACGCGC	635
Db	2124	GGAAGCACCGCGGAGGAGCATCGAAGAAACGTCCACGSGACTGTGTGATGCGCGCGCTCAAGG	2183
QY	636	CTCGATGTTCAACAAG	652
Db	2184	CCCCACGCGCAGCTGG	2200

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd..

OM nucleic - nucleic search, using sw model

Run on: April 22, 2003, 01:28:39 ; Search time 1961 Seconds
(without alignments)
9761.890 Million cell updates/sec

Title: US-09-673-198-31

Perfect score: 1182
Sequence: 1 atgcgcagcctgtcgcattt.....gtctcgcagcagcaaaag 1182

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

EST:*

1: em_estba:**
2: em_esthum:**
3: em_estlin:**
4: em_estnu:**
5: em_estov:**
6: em_estpl:**
7: em_estro:**
8: em_hlc:**
9: gb_est1:**
10: gb_est2:**
11: gb_hlc:**
12: gb_est3:**
13: gb_est4:**
14: gb_est5:**
15: em_estfun:**
16: em_estom:**
17: gb_gss:**
18: em_gss_hum:**
19: em_gss_inv:**
20: em_gss_pln:**
21: em_gss_vrt:**
22: em_gss_fun:**
23: em_gss_mam:**
24: em_gss_mus:**
25: em_gss_other:**
26: em_gss_pro:**
27: em_gss_rtd:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	205	17.3	570	13	B1721015
2	153.8	13.0	280	17	A2577432
3	129.4	10.9	687	10	BE213317
4	127.4	10.8	488	17	BH001147
5	124.4	10.5	626	10	BE519766
6	118.4	10.0	680	13	B471224

C	7	113.8	9.6	784	13	B1751397
	8	112	9.5	1838	11	AV109535
	9	110.6	9.4	446	12	BF428515
	10	109.4	9.3	543	13	BM406343
	11	109.4	9.3	594	14	BO518501
	12	107.4	9.1	687	10	AM775985
	13	105	8.9	838	12	BG588663
	14	98.8	8.4	798	10	BE412040
	15	97.2	8.2	712	12	BG586951
	16	93.8	7.9	619	13	B1470380
	17	91.6	7.7	469	13	BI894320
	18	91.6	7.7	546	13	BI674464
C	19	91	7.7	706	17	CNS010LS
	20	90.2	7.6	584	13	BI942645
	21	90.2	7.6	845	13	BF261910
	22	89.6	7.6	684	10	BG840487
	23	88.8	7.5	831	12	BG343005
	24	88.6	7.5	598	10	AV934537
	25	86.4	7.3	471	10	AM616251
	26	86	7.3	572	13	BI943128
	27	85.6	7.2	575	13	BI943466
	28	85	7.2	548	14	BQ039624
	29	84.6	7.2	594	10	AV933746
	30	84.6	7.2	829	13	BM112204
	31	82.6	7.0	751	12	BF005132
	32	82	6.9	468	13	BI923028
	33	81	6.9	591	9	AT993895
	34	81	6.9	637	12	BG043747
	35	80.8	6.8	333	9	AL819484
	36	80.4	6.8	555	13	BM524138
C	37	80	6.8	925	17	CNS0091P
	38	79.8	6.8	675	12	BG449020
C	39	79.4	6.7	935	17	CNS006XK
	40	79.2	6.7	286	10	AM067754
	41	78.8	6.7	562	13	BI472011
	42	78.8	6.7	596	14	BQ295914
	43	78.8	6.7	670	12	BG453943
	44	78.8	6.7	692	12	BF598623
	45	78.8	6.7	743	14	BQ917023

ALIGNMENTS

RESULT 1
LOCUS B1721015 570 bp mRNA linear EST 19-SEP-2001
DEFINITION 1031053E04.y1 C. reinhardtii CC-1690, Stress II (normalized),
Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION B1721015
VERSION B1721015.1 GI:15696710
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE 1 (bases 1 to 570)
AUTHORS Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre,
P., McDermott, J.E., Shrago, J., Sillow, C., and Stern, D.
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1031
JOURNAL Unpublished (2001)
COMMENT Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu

FEATURES

FEATURES
source Location/Qualifiers
1..570
/organism="Chlamydomonas reinhardtii"

```

/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized)
), Lambda zap II"
/note="Vector: Bluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffrey McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)
and shifted to TAP - NO3- (24hrs); H2 production
conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
zap II (Stratagene) in the EcoRI (5') and XhoI (3')
sites. Bluescript II SK- plasmids were excised from the
lambda zap clones by superinfection with ExSist
(Stratagene) phage. The library was normalized using
method 4 described in Bonaldi et al., (1996) genome
Research 6: 791-806."
BASE COUNT      108 a      211 c      161 g      87 t      3 others
ORIGIN

```

```

Query Match      17.3%; Score 205; DB 13; Length 570;
Best Local Similarity 63.7%; Pred. No. 1.4e-27;
Matches 310; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

QY 366 CAACAGAGAAAGCCTCTGACGGCAGCACTCTGATGCGAGCGCCAGAGAAAGC 425
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2 CACGAGGAGAACCTCATGTCGGGCGGCCCTTCATCTCGCGCTGCCAGAGTAGCG 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 426 CGGCACATCTCGCGGTGACACGACGACCTCGCGCTTTAGAGCGCTGGCGGCGA 485
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 62 CTCGAATCTCTGCGCCCGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 486 GGAACAGCGCTGCGTCAGCGCGCTCATCATCAGCGCGCTCGCGCGCGCTCGCG 545
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 122 GCCCGAGGCGCGCTGCGCGCGCTCATCTCAGCGCTCGCGCGCGCTCGCGCG 181
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 546 GAGCCTGAGCGCATCGCGCGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCG 605
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 182 GCCCGTGAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 241
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 606 CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 665
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 242 CATTGGGCAAGATCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 301
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 666 CGAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 725
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 302 TGAAGCTCACTACTCTGCGAGTGGAGTACGACACATGATGATGATGATGATG 361
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 726 ATCCATGTCATGCGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 785
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 362 GTCCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 421
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 786 CGCGGACATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 845
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 422 GCCCGAGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 481
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 846 CGTCGCC 852
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 482 GGAGGCC 488
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 2
A2577432/c      280 bp      DNA      linear      GSS 08-DEC-2000
LOCUS           10a03 Shot-gun genomic library of Rhizobium strain ANU265 Rhizobium
DEFINITION      sp. NGR234 genomic clone 10a03, DNA sequence.
ACCESSION       A2577432
VERSION         A2577432.1 GI:11603701
KEYWORDS        GSS
SOURCE          Rhizobium sp. NGR234.

```

```

ORGANISM        Rhizobium sp. NGR234
REFERENCE       Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
                Rhizobiaceae; Rhizobium.
AUTHORS        Vipro, V., Rosenthal, A., Broughton, W.J. and Perret, X.
TITLE          Genetic snapshots of the Rhizobium species NGR234 genome
JOURNAL        Genome Biol. 1 (6), RESEARCH0014 (2000)
MEDLINE        2114532
COMMENT        Contact: Virginie Vipro
                Laboratoire de Biologie Moléculaire des Plantes Supérieures
                Université de Genève
                1 Chemin de l'Imperatrice, Chambesey/Genève 1292, Switzerland
                Tel: +44(0)1603450000
                Fax: +44(0)1603450045
                Email: virginie.vipro@obs.jc.ac.uk
                Class: shotgun.
FEATURES        location/Qualifiers
SOURCE          1. 280
                /organism="Rhizobium sp. NGR234"
                /strain="ANU265"
                /db_xref="taxon:394"
                /clone_lib="10a03"
                /clone_lib="Shot-gun genomic library of Rhizobium strain
                ANU265"
                /note="Vector: M13; derivative strain of NGR234 cured of
                pNGR234a"
BASE COUNT      53 a      83 c      91 g      53 t
ORIGIN

```

```

Query Match      13.0%; Score 153.8; DB 17; Length 280;
Best Local Similarity 73.2%; Pred. No. 2.4e-18;
Matches 197; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 571 ACCGTGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 630
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 271 AGCGCGCAATAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 212
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 631 AGCGCTCGATGTTCAACAGCGCGCTGAGCTGATGACAGCGCGCGCGCGCGCG 690
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 211 AGCGCTCGATGTTCAACAGCGCGCTGAGCTGATGACAGCGCGCGCGCGCGCG 152
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 691 GAGCGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 750
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 151 CGCCCGCAACGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 92
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 751 TTCTGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 810
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 91 TACAGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 32
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 811 TTGCGCTGACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 839
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 31 TATGCGCTCACTATCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 3
BE213317/c      687 bp      mRNA      linear      EST 30-JUN-2000
LOCUS           BE213317
DEFINITION      EST0074 Triticum aestivum Lambda zap Triticum aestivum cDNA clone
                JAL1D.H11.T3.5' similar to putative 1-deoxy-D-xylulose-5-phosphate
                reductoisomerase, mRNA sequence.
ACCESSION       BE213317
VERSION         BE213317.1 GI:8845070
KEYWORDS        EST.
SOURCE          Triticum aestivum
ORGANISM        Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
                Triticeae; Triticum.
REFERENCE       Anderson, J.M., Williams, C.E. and Goodwin, S.B.
                Analysis of an EST database reveals a probable CF2 resistance gene
                homolog in wheat
                unpublished (2000)
JOURNAL

```


FEATURES						
source		Location/Qualifiers				
		1..680				
		/organism="Hordeum vulgare subsp. vulgare"				
		/cultivar="Haruna Nijo"				
		/db_xref="taxon:112509"				
		/clone="baallf12"				
		Nijo adult, heading stage top three leaves"				
		/dev_stage="adult, heading stage"				
		/issue_type="top three leaves"				
BASE COUNT	180 a	145 c	181 g	174 t		
ORIGIN						
Query Match	10.0%	Score 118.4;	DB 13;	Length 680;		
Best Local Similarity	49.9%;	Pred. No. 5.8e-12;				
Matches 333;	Conservative	0;	Mismatches 326;	Indels 9;	Gaps 1;	
OY	79	CCCCGAGCGTTCCGCACCCGCTGCTTGACACGGCGGGGCCAACATCCGGCGACTGGCCGAA	138			
Db	9	CCTGATTAAGTTCGCGGGTTGTGTGCTCTACTGCTGGCTCCAAGTCACTTCTTACTGAT	68			
OY	139	ATGGGGCGCTGGCGTGAAGGCGGAGCTTCCGCGACCGGGCATGAGAGACTGCTGGCCGG	198			
Db	69	CAGGTGAACAACGTTCAAACCAAGCTGGTTGGCTTAAGGAGCAGATCTATTCTTAACAG	128			
OY	199	CTGGCGGAGGGCGTGGCGG-----GACGGGACCGAGGTGCGGGCGGGCGCGAG	249			
Db	129	CTAAGAAGCGGTTAGTGTGTGTGAGAGATCCGGAAATTAATTCCTGGGAGCAAGT	188			
OY	250	GCATTCGCCGAGGCGCGCGACCGCCGCGCCGACCTGGACATCTGGCCATCTGGGCGCC	309			
Db	189	GTCATAAGAGGTTGCTGCGCACCCAGATSCAGTTACAGTGTTCGCGGATTAAGTGGTGT	248			
OY	310	GCGGGCTCTGTCGCCGCGGATGGGGCGGTGAAGCAGCGCGGACGCTGGCGCTGCCAAC	369			
Db	249	GCAGGACTCAAGGCTACAGAGTTGCACCAGTTGAAGCTGGGAAGACATAGCACTGGCGAAC	308			
OY	370	AAGGAAAAGCTCTGTGACGGCAGGGCACTCCTGATGCGGACGCGCCAGAGAAAGGCGCC	429			
Db	309	AAGGAGACACTTAATTGACAGGGGCGCGTTCGCTTCCTCCCTTGCGCACAGCAATGTG	368			
OY	430	ACGATCTCGCGGTGGACAGCGAGCACTCCGGGGCTTTCAGGCGCTGGCGGGCGAGAC	489			
Db	369	AAAAATCTTCTGCTGATTCACAGACACTCTGCAATATTTCAGTGTATACAGAGCTTGTCT	428			
OY	490	ACGGCCTCGTCGACGCGGCGTCAATCAACGGCGTCCGGCGGGCGCTTCGCGACTGGAGC	549			
Db	429	GAAAGATCACTTCGTCGCTTATCTCTGACTGGCTGGCGGGTTCAGGAGACTGGCCA	488			
OY	550	CTCGGCGCATTCGCGGCGCTGCACAGCTGGCGGCGGAGGCGCCATCCCACTGTCTCAAG	609			
Db	489	GTAGAAGAGCTGAAGAAGAGCTAAGAGTTGCGCGAGTCTTTGAAAGCACCCAAATTTGAGCATG	548			
OY	610	GCGCAGCGGATCTCCATCGACAGCGCTTCGATGTTTCAACAMAGGCGCTCGAGCTGATCGAG	669			
Db	549	GGGAAGAAATACACAGTAGATTTCTGCTACTTATTATTCACAAGAAGGGTGAAGATTTCGAG	608			
OY	670	ACGGCGCAATTCCTGGCTTCGAGCGCGGACCGGAGCTGGAGCGGTGCTCATCCGAATCC	729			
Db	609	GCACATTTATTTGTTGTGGCTGTAATATATGACATTGACATTGTATTCACCCCGAATCC	668			
OY	730	ATCGTCCA 737				
Db	669	ATCATACA 676				
RESULT 7						
BI751397/c	BI751397	784 bp	mRNA	linear	EST 25-SEP-2001	
LOCUS	Ta01_20B04_R					
DEFINITION	Ta01_AAFc_EBCRC_Fusarium_graminearum_inoculated_wheat_heads					
	Triticum aestivum cDNA clone Ta01_20B04, mRNA sequence.					
ACCESSION	BI751397					

VERSION	BI751397.1	GI:15773199
KEYWORDS	EST.	
SOURCE	Bread wheat.	
ORGANISM	Triticum aestivum	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae ; Triticeae; Triticum. 1 (bases 1 to 784)	
AUTHORS	Ouellet,T., Dan,H., Koull,A., Chappados,J., Couroux,P., De Moors,A., Harris,L.J., Hattori,D.I., Robert,L.S., Singh,J.A., Sprout,D. and Tinker,N.A.	
JOURNAL	Expressed Sequence Tags from Wheat Heads 24 Hours after Spray Inoculation with Fusarium graminearum	
COMMENT	Contact : Ouellet, Therese Eastern Cereal and Oilseed Research Centre Agriculture and Agri-Food Canada Neelby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A 0C6, CANADA Tel: (613) 759-1658 Fax: (613) 759-1701 Email: ouellet@erem.agr.ca.	
FEATURES	Location/Qualifiers	
SOURCE	1..784	
	/organism="Triticum aestivum"	
	/cultivar="Frontana"	
	/db_xref="taxon:4565"	
	/clone="Ta01.20b04"	
	/clone_lib="Ta01.ARC_ECORC_Fusarium_graminearum_inoculate d_wheat_heads" /tissue_type="heads" /dev_stage="anthesis" /note="Vector: pCEM-T easy; site_1: EcoRI; site_2: EcoRI; Controlled chamber-grown wheat heads were spray inoculated at mid-anthesis with a Fusarium graminearum microconidia suspension (50,000 spores/ml) and kept under intermittent mistling for 24 hours, then collected and immediately frozen in liquid nitrogen."	
BASE COUNT	172 a 237 c 187 g 185 t	3 others
ORIGIN		
Query Match	9.6% Score 113.8; DB 13; Length 784;	
Best Local Similarity	53.2%; Pred. No. 3.9e-11;	
Matches 292; Conservative	0; Mismatches 247; Indels 10; Gaps 2;	
Dy	522 GTCCGGCGGCGCGCTTCGCCAGACTGGAGCGCATCCGCGCCTGCACCGTGCCGA	581
Db	784 GTCTGGCGGTCTTTCAAGAGCACTAGACCAGTTGAGAAGCTGAAAGATGTAAGGTTGCCGA	725
Dy	582 GGCGCAGCGCCATCCCACTGTCATCGGCGCAGCGGATCTTCACCGCACAGCGCTGCAT	641
Db	724 TGCTTTGAAGAACCCAACTGGAGCATGGGGAAGAAAATACAGTAGATTCTGCTACTTT	665
Dy	642 GTTC-AACAAGCGCGTCAGCTGATCGAGAGCGCGGAATTTCTGGCTTCGAGCGGACG	700
Db	664 GTTCAAACAAGGGGTAGAGATTATCGAGGGCGCATATTGTGGTGTCTGTAATGTATG	605
Dy	701 GGATCGAGGCGGTGTCATCCGCAATCTCATCTGTCATGCGATGTGGGCTTTCGACG	760
Db	604 ACATTGAAATTTGTGATTCACCCACAGTCCATCATACACTATGATTTAAAAACCCAGGATT	545
Dy	761 GGGGCGTATGAGCCCATTCGAGCCCCCGCGCAGATGCGCGCAGCCATCGGATTCGGCGGA	820
Db	544 CATCTGTCTTGAGCTACGCTGGGATGGCCAGACATGCGGCTACCCGATCTTATACACTTGT	485
Dy	821 ACTGCGCGCGT-----GCGGCGAGGTGCCCGTGCACCGGATTCGACTCGCACAGA	871
Db	484 CTTGCGCAGACGAGGTACTAGTCTCGAGGTCCAGCTCGGCCCGCGCTTAGACCTTTGCAAGC	425
Dy	872 TTGCGAGCCTTACCTTCAGAAAGCTGACGAGAAAGCTTTTCGGCCCTGAGGCTTGGCG	931
Db	424 TGGGTTCCTGACATTTAAAGCTCCCGCAACGTGAATAATACCATTCGGGTGATCTCGCGT	365

QY 932 GAGACGTCATGCGCGCGCGCTGTGCGGCGCCCTTCACGCGCGCAAGAGATCG 991
 Db 364 ATGCGGCGAGGCGCGCGCTGACACCATGAGCGAGTTTGACGCGCTCAACAGAGAG 305
 QY 992 CGCTCGATCATTTATCCCGGAGCGCATCGGCTTTCTGACATGCGCGCGTGTGCGAG 1051
 Db 304 CGGCGAGCTTTATTCAGCAAGAAAGATGACTACCTGACATCTTCAGAGTGTGAGAGA 245
 QY 1052 AGACGCTCG 1060
 Db 244 TGACGTGCG 236

RESULT 8
 LOCUS AY109535 1838 bp mRNA Linear HTC 25-MAY-2002
 DEFINITION Zea mays CL389_1 mRNA sequence.
 ACCESSION AY109535
 VERSION AY109535.1 GI:21213292
 KEYWORDS HTC.
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 1838)
 Hainey,C.F., DoJian,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
 Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
 Maize Mapping Project/Dupont Consensus Sequences for Design of
 Overgo Probes
 Unpublished (2002)
 2 (bases 1 to 1838)
 Coe,E.C.
 Direct Submission
 Submitted (25-APR-2002) Maize Mapping Project, University of
 Missouri, Columbia, MO 65211, USA

FEATURES
 source
 Location/Qualifiers
 1..1838
 /organism="Zea mays"
 /db_xref="MaizeDB:631829"
 /db_xref="taxon:4577"
 /clone="CL389.1"
 /clone_ltb="Maize Mapping Project/Dupont Consensus
 Library"
 /note="this sequence is part of a project of EST
 assemblies resulting from the application of public
 contigs to seed Dupont contigs; this resource was
 assembled by Dupont as part of a collaboration for the
 overgo addressing of BACS in conjunction with the Maize
 Mapping Project"

BASE COUNT 438 a 398 c 433 g 438 t 131 others
 ORIGIN

Query Match 9.5% Score 112; DB 11; Length 1838;
 Best Local Similarity 44.9%; Pred.No.8.3e-11;
 Matches 347; Conservative 0; Mismatches 416; Indels 9; Gaps 1;

QY 293 CGGCATCTGTGGCGCGCGCTGTGCGGCGCGCGATGCGCGCGCTGAAGCAGCGCGCA 352
 Db 707 CAGGATAGTAGTTGTCAGGCGCTGAAGCCTGACGTTGCGCAATTGAACCTGTGAAG 766
 QY 353 CGCTGGCGCTGCGCAACAGGAAGCCTGTCGACGCGCGCAACTCTGTATGCGGACGG 412
 Db 767 ACATAGATTTGGCAACAGAGACACTTATTGACAGGTGCTTTGCTTCCCTTGG 826
 QY 413 CCCAGGAGAGCGCGCGCGCTGCTGCGCGCGCGAGCAGCAGCAGCAGCGCGCTTTTCAG 472
 Db 827 CACCAACACCAAGTAGTAATTTCTTCAGCTGATTTCTGACACCTCTGNNNNNNNNNN 886
 QY 473 CGCTGGCGCGCGAGACACGCGCTGCTGAGCGCGCTATCATCAGCGCGCTCGCGCGCG 532
 Db 887 NNN 946

QY 533 CGTTCGAGCTGAGACCTCGAGCGCATCGCGCTGACCGCGCGAGCGAGCGCC 592
 Db 947 CTTTCAGGAGCTGGCGCACTTTCAGAGCGCTGAAGAGATGTAAAGTTGCTGACCGCTTAAAGC 1006
 QY 593 ATCCCACTGCTGATGGCGCGAGATCTCTCATCGACAGCGCTCGATTTTCAACAAG 652
 Db 1007 ATCCAACTGGAATATGGGAAGAGATACAGATGATTTCTTCACTTATTTCAACAAG 1066
 QY 653 CGCTGAGCTGATCGAGAGCGCGCGCAATCTTGGCTTGCAGCGCGAGCGAGCGCGG 712
 Db 1067 GTTTAGAGTTATTTGAGACACATTTATTTATTTGCTGAAATATGATGATTTGATG 1126
 QY 713 TCGTCATCCGCAATTCATGCTCATGCGATGATGGCGCTTGCAGCGCGCGCTGATG 772
 Db 1127 TGATTCACCCACAGCTCTCATACCTATATGTTGAACCCAGAGATTCATCTGCTAG 1186
 QY 773 CCCATCTGCGCGCGCGCGCGCATGCGCGCATGCGATGATTTGGCGCTGACCTGCGCGG 829
 Db 1187 CTCAGTTGGATGGCCACATATGCGGTTACCAATCTTATACCTTATCATGCGCGCAT 1246
 QY 830 -----GTGCGCGGAGGTGCGCGCGCGCGATCGACCTCGACAGATTCGAGCGCTCA 883
 Db 1247 GAATCTATGCTCTGAGAGTACCTGCGCGCGCTGATCTTTGCAAGTTGGTTCACTGA 1306
 QY 884 CTTTCAGAACCTTGACGAGGAGCGCTTTCGCGCGCTGAGCGCTTGCAGACGTCATG 943
 Db 1307 CATTGAGAGCTTCAGACAGCAACCTAAATATCCATCAATGACCTGATGACCTGAGCC 1366
 QY 944 CGCGCGCGCGCTGTGCGCGCGCGCGCTTCACAGCGCGCGCAAGAGATGCGCGCTCATCT 1003
 Db 1367 GCGCTNNNNNCACATGACAGAGTCTGAGCGCGCTGTAAGAGAGCGCGTGAAGTTGT 1426
 QY 1004 TCATCGCGGAGCGCGCGCTTTCGACATGCGCGCGCGCTGAGAGAGAC 1055
 Db 1427 TCATTGAGAGAAAGATTAGCTACCTGACATATTCAGATGAGTGTGAGCTTAC 1478

RESULT 9
 LOCUS BF428515 446 bp mRNA Linear EST 29-NOV-2000
 DEFINITION WHE1409_A03_A056S wheat drought stressed leaf cDNA library Trilicium
 aestivum cDNA clone WHE1409_A03_A05, mRNA sequence.
 ACCESSION BF428515
 VERSION BF428515.1 GI:11440279
 KEYWORDS EST.
 SOURCE bread wheat.
 ORGANISM Trilicium aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
 1 (bases 1 to 446)
 Trilicium; Trilicium.
 Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Han,P.S., Hsiao,C.C.,
 Kang,Y., Iazo,G.R., Miller,R., Nguyen H.T., Rausch,C.J., Seaton
 C.L., Tong,J.C. and Zhang,D.
 The structure and function of the expressed portion of the wheat
 genomes - Drought stressed leaf cDNA library
 Unpublished (2000)
 Contact: Olin Anderson
 US Department of Agriculture, Agriculture Research Service, Pacific
 West Area, Western Regional Research Center
 800 Buchanan Street, Albany, CA 94710, USA
 Tel: 5105959773
 Fax: 5105959818
 Email: oanderson@wr.usda.gov
 Sequence have been trimmed to remove vector sequence and low
 quality sequence with phred score less than 20
 Seq primer: Stratagene SK primer.
 Location/Qualifiers
 1..446
 /organism="Trilicium aestivum"
 /cultivar="TAM W101"
 /db_xref="taxon:4565"
 /clone="WHE1409_A03_A05"

COMMENT

Contact: Research Genetics, Libraries Division
Tel.: 1-800-711-6195
Email: cdna@resgen.com
For clone info: please contact Research Genetics, Libraries Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: T3.

FEATURES

source	location/Qualifiers
	1..543
	/organism="Solanium tuberosum"
	/cultivar="Kennedec"
	/db_xref="taxon:4113"
	/clone="cPRO2706"
	/clone_1id="potato roots"
	/tissue_type="roots"
	/dev_stage="in vitro grown stem cuttings"
	/lab_host="SDLR"
	/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Corneil University, Tanksley lab; sequencing: The Institute for Genomic Research. Roots were isolated from in vitro grown stem cuttings on CM medium. Roots were isolated two weeks after placing the stem cuttings from in vitro grown plants on medium."

BASE COUNT

148 a	119 c	121 g	155 t
-------	-------	-------	-------

ORIGIN

```

Query Match          9.3%; Score 109.4; DB 13; Length 543;
Best Local Similarity 54.0%; Pred. No. 2.4e-10;
Matches 224; Conservative 0; Mismatches 191; Indels 0; Gaps 0;
```

OY	415	CAGGAGAAGCGGGCCACATCTCTGCCGTGCAGACGACGACACTCCGGGTCCTTCAAGCGC	474
Db	2	CACAAGCATTAAGTAGAAGATTCTCTCCAATTCAGAACTACGCTATATATTCACATGTC	61
OY	475	CTGGCGGCGAGAGACACGCGCTGCGTGCGAGCGCTCATCATCACGCGCGCTCCGCGCGCGC	534
Db	62	ATAACAAGCTTTGCCAGAGAGGTGCGCTTAGGGGCAATTAATATACATGCACTGSGAGGGGCT	121
OY	535	TTCGCGCACTGGAAGCCTTGAGCGGCAATCCGCGCTGACCCTGGGCGGAGCGCAGGCCCAT	594
Db	122	TTTCAGGGGACTGGCCAGTTGAGAAAGTTGAAGAAGTTAAAGTAGCTGATCTTTGAAGCAT	181
OY	595	CCCACATGCTCATAGGGGCGAGGGGATCTCCATTCGACACGCGCTCGATGTTCACAAAGCG	654
Db	182	CCCCATTGGAACATBGGAAAAAGATTTACTGTGATTTCTGCCACATTTATTCATATAAGGT	241
OY	655	CTCGAGCTGATCGAGACGCGGCAATTTCTTGCGCTTCGAGCGGAGCCGAGATCGAGCGGCT	714
Db	242	CTTGAAATTATCGAAGCTACACTACTTTTCGAGCGGAGTATGACACAACTTGAATTTGTC	301
OY	715	GTCGATCGGCAATCATGTGTCATGCCATGGTGGGCTCTCTCGACGAGGGGGCTATGGCC	774
Db	302	ATTCATCCCAATCCATCATATCATTTGATGGTGTGAACACAGAGATTCATCAATATTTGCA	361
OY	775	CATCTGCGCGCCGCGCATGCGCCACGCAACATTCGCGCTCAACTGGCGCG	829
Db	362	CACTAGSGGTGGCTGATATGCGTTTGCCCACTTTTATCTTATCTCTGGCGCG	416

RESULT 11

LOCUS	BO518501	594 bp	mRNA	linear	EST 10-JUN-2002
DEFINITION	EST625916 Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues Solanum tuberosum cDNA clone STMK54				
ACCESSION	BO518501				
VERSION	BO518501.1 GI:21377370				
KEYWORDS	EST.				
SOURCE	Solanum tuberosum				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum.				

REFERENCE 1 (bases 1 to 594)

/organism="Medicago truncatula"
 /cultivar="A17"
 /db_xref="taxon:3880"
 /clone="PMHRP-57K17"
 /clone_lib="MHRP-"
 /tissue_type="roots"
 /dev_stage="phosphate-starved"
 /lab_host="X10LR
 /note="Vector: palnucscript SK-; Site_1: EcoRI; Site_2:
 XhoI; At the trifoliolate stage, M. truncatula plants were
 transplanted to phosphate-free sand and grown for a
 further 30 days. During this period, they were fertilized
 twice weekly with 1/2 Hoaglands solutions containing 20mM
 potassium phosphate. cDNA was prepared from polyA+
 enriched RNA. The cDNA was directionally ligated into the
 Uniapack XR vector from Stratagene and packaged using
 GigaPack III Gold packaging extracts. Plasmids containing
 cDNA inserts were excised from the recombinant lambda-Zap
 phage using Ex-assist helper phage and propagated in
 X10LR cells."

AUTHORS
Anderson, U. A., Appels, R., Bailey, P., Blake, T., Closter, T., Cloutier,
S., Dubcovsky, J., Feuille, C., Gale, M., Graner, A., Gustafson, P.,
Herrmann, R. G., Holton, T., Jacquemin, J. M., Jia, J., Joudrier, P.,

Db 421 CTGGCCCTGGCTTGATCTTTGCAGCTTGTTCCCTTACTTTAAAGCTCTGACACA 480
QY 905 AACGCTTTCGGCCCTGAGGCTTGGCGAGAGCTCATGGCGGCGGCGCTGTGGGGC 964
Db 481 TTAGTATCCATCATGAATCTTGCTATGCTGTGGCGGCTGAGGGCACCATGACAG 540
QY 965 CCGCTTTCACGGCGGCAAGAGATGCGCTCGATCATTTTCATCGCCGAGCGCATCGGGT 1024
Db 541 GTGTCTTAGTGCAGCAATGAGAAAGCTGTAGAGATGTTTATGATGAAGAAAGATTAGCT 600
QY 1025 TTCTGGACAT 1034
Db 601 ATTGGATAT 610

Search completed: April 22, 2003, 03:27:01
Job time : 1985 secs

Methylinomas 16a S
Methylinomas 16a O
Plasmid_P1US75.0xS
Neisseria meningit
Mycobacterium tube
Mycobacterium tube
Neisseria meningit
Neisseria gonorrh
Isoprenoid related
1-deoxy-D-xylose
Propionibacterium
Isoprenoid related
Bacillus lichenifo
Isoprenoid related
C glutamicum codin
C glutamicum codin
C glutamicum codin
Polynucleotide seq
1-deoxy-D-xylose
Listeria monococ
Listeria innocua c
A. thaliana 1-deox
A. thaliana DXPRI
Isoprenoid related
1-deoxy-D-xylose
Corynebacterium gl
Arabidopsis herbic
1-deoxy-D-xylose
Arabidopsis herbic
1-deoxy-D-xylose
Isoprenoid related
DNA encoding novel
Chlamydia pneumoni
Nucleotide sequenc
Arabidopsis thalia

XX PA (KYOW) KYOMA HAKKO KOGYO KK.
 XX PA
 XX PI Miyake K, Hashimoto S, Motoyama H, Ozaki A, Seto H, Kuzuyama T,
 XX PI Takahashi S;
 XX DR WPI: 1999-620434/53.
 XX P-PSDB: AAY52839.
 XX
 XX Preparation of recombinant isoprenoid compounds useful for treatment of
 XX heart diseases, osteoporosis and hemostasis, preventing cancer and
 XX immunopotentialation
 PS Claim 7; Page 133-136; 145pp; Japanese.
 XX
 CC The present invention describes the preparation of an isoprenoid
 CC compound comprising using at least 1 DNA e.g. encoding proteins which
 CC elevate the efficiency of the synthesis or DNA encoding a farnesyl
 CC pyrolic producing enzyme. The method of preparation of an isoprenoid
 CC compound comprises using at least 1 DNA, a vector, cloned cells, their
 CC derived recombinant DNAs or transformed products in a culture system and
 CC extracting the isoprenoid accumulated in the medium. The DNA encodes at
 CC least 1 of the following: (a) a compound for activating or catalysing
 CC the production of 1-deoxy-D-xylulose-5-phosphate from pyruvic acid and
 CC glyceraldehyde-3-phosphate; (b) an enzyme producing farnesyl pyrolic
 CC acid; (c) a protein which elevates the efficiency of synthesis of
 CC isoprenoid compounds and comprises a 3 or 4 amino acid sequence
 CC optionally with 1 or more of the amino acids being deleted or
 CC substituted with an additional amino acid being inserted; (d) a protein
 CC which activates or catalyses the production of 2-C-methyl-D-erythreitol-
 CC 4-phosphate from 1-deoxy-D-xylulose-5-phosphate; or (e) a protein which
 CC activates a target compound or reaction and is a string end or hybrid of
 CC the DNA encoded in (a)-(e). Isoprenoid compounds are useful in drugs
 CC (e.g. for the treatment of heart diseases, osteoporosis and hemostasis,
 CC for preventing cancer and as immunopotentialators), health foods and
 CC antifouling coatings. The isoprenoids also inhibit enzymatic reactions
 CC on the non-nevonal pathway and can be used as antibacterials and
 CC herbicides. The present sequence is used in the exemplification of
 CC the present invention.
 XX
 XX Sequence 1182 BP; 185 A; 408 C; 417 G; 172 T; 0 other;
 SO
 Query Match 100.0%; Score 1182; DB 20; Length 1182;
 Best Local Similarity 100.0%; Pred. No. 1.7e-187;
 Matches 1182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 421 AACGGCCACGATCTCCGCTGGACAGCAGCACTCCGCGTCTTACAGCCCTGGCG 480
 DB 421 AACGGCCACGATCTCCGCTGGACAGCAGCACTCCGCGTCTTACAGCCCTGGCG 480
 QY 481 GGGCAGGACAGCGGCTGGCTGAGCGGCTCAATCAACAGCGGCTCCGCGGCGCTTCCG 540
 DB 481 GGGCAGGACAGCGGCTGGCTGAGCGGCTCAATCAACAGCGGCTCCGCGGCGCTTCCG 540
 QY 541 GACTGAGACCTTCGAGCGCATCCGCGCTGCACCTGGCGAGGCGGCGCATCCCAAC 600
 DB 541 GACTGAGACCTTCGAGCGCATCCGCGCTGCACCTGGCGAGGCGGCGCATCCCAAC 600
 QY 601 TGTTCATGAGGCGCAGCGGATCTCATGACAGCGCTCATGATGTTCAACAGCGCTCGAG 660
 DB 601 TGTTCATGAGGCGCAGCGGATCTCATGACAGCGCTCATGATGTTCAACAGCGCTCGAG 660
 QY 661 CTGATCGAGACGCGCGAATTTTGGGCTTCAGCGCGGACCGGATCGAGGGGTGTCAT 720
 DB 661 CTGATCGAGACGCGCGAATTTTGGGCTTCAGCGCGGACCGGATCGAGGGGTGTCAT 720
 QY 721 CCGCAATCATCGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 DB 721 CCGCAATCATCGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 QY 781 GGGCCCGCGACATGCGCCACGCGCATGCGGATTCGCTGAACCTGCGCGCGCGAG 840
 DB 781 GGGCCCGCGACATGCGCCACGCGCATGCGGATTCGCTGAACCTGCGCGCGCGAG 840
 QY 841 GTGCGCGCGCGCGCGATGACCTGCGACAGATTCGCGAGGCTTCCAGAACCTGAC 900
 DB 841 GTGCGCGCGCGCGCGATGACCTGCGACAGATTCGCGAGGCTTCCAGAACCTGAC 900
 QY 901 GAGGAGCGCTTCCGCGCTGAGGCTTCCGCGAGGCTTCCGCGAGGCTTCCGCGAGG 960
 DB 901 GAGGAGCGCTTCCGCGCTGAGGCTTCCGCGAGGCTTCCGCGAGGCTTCCGCGAGG 960
 QY 961 GGGCGCGCTTCAACGCGCGCGAGAGATGCGCTCATCTTTCATGCGCGAGCATC 1020
 DB 961 GGGCGCGCTTCAACGCGCGCGAGAGATGCGCTCATCTTTCATGCGCGAGCATC 1020
 QY 1021 GGGTTTCGACATGAGGCGGCGGTGTCGAGAGAGCGTCCGCGGCTTTCGACGAGCC 1080
 DB 1021 GGGTTTCGACATGAGGCGGCGGTGTCGAGAGAGCGTCCGCGGCTTTCGACGAGCC 1080
 QY 1081 CTGTTTCGAAAGTGCAGCGCTTTCGAGAGAGTCTGCGCATGTCCTGCTCG 1140
 DB 1081 CTGTTTCGAAAGTGCAGCGCTTTCGAGAGAGTCTGCGCATGTCCTGCTCG 1140
 QY 1141 AGAGCGGAGAGAGAGCGCGGCTTCCGCGAGGCTTCCGCGAGGCTTCCGCGAG 1182
 DB 1141 AGAGCGGAGAGAGAGCGCGGCTTCCGCGAGGCTTCCGCGAGGCTTCCGCGAG 1182

RESULT 2
 AAL40184
 ID AAL40184 standard; DNA; 1161 BP.
 XX
 AC AAL40184;
 XX
 DT 13-SEP-2002 (first entry)
 XX
 DE Isoprenoid related nucleic acid sequence SEQ ID No 96.
 XX
 KM Isoprenoid; CoQ(10); 1-deoxyxylulose-5-phosphate synthase; DXS; DDS;
 XX decaprenyl diphosphate synthase; ds.
 OS Sphingomonas trueperi.
 XX
 PN WO200226933-A2.
 XX
 PD 04-APR-2002.
 XX

PF 28-SEP-2001; 2001WO-US30328.
 XX
 XX 29-SEP-2000; 2000US-236580P.
 XX
 XX (CRGI) CARGILL INC.
 XX
 XX Gokarn R, Jessen H, Zidwick MJ;
 XX
 XX MPI: 2002-416480/44.
 XX
 PT Substantially pure polypeptides having e.g.,
 PT 1-deoxyxylulose-5-phosphate synthase activity, useful for the
 PT production of isoprenoids, especially Coq(10)
 XX
 PS Claim 48; Fig 29; 246pp; English.
 XX

CC The invention relates to methods and materials for the production of
 CC isoprenoids. More particularly the invention provides isolated nucleic
 CC acids, substantially pure polypeptides, host cells, and methods for
 CC producing various isoprenoid compounds. The polypeptides are useful for
 CC the production of isoprenoid compounds, especially Coq(10). Expressing the pure
 CC polypeptides, which has 1-deoxyxylulose-5-phosphate synthase (Dxs)
 CC activity or decaprenyl diphosphate synthase (Dds) activity, is useful for
 CC increasing production of Coq(10) in a cell having endogenous Dds
 CC activity. This polynucleotide represents a nucleic acid sequence relating
 CC to the isoprenoid production of the invention.
 XX

Sequence 1161 BP; 178 A; 370 C; 416 G; 197 T; 0 other;

Query Match 32.2%; Score 380.4; DB 24; Length 1161;
 Best local similarity 60.3%; Pred. No. 1.1e-54;
 Matches 649; Conservative 0; Mismatches 421; Indels 6; Gaps 1;

OY 8 GCCTGCGATCTTTGGGGCCACCGGCTCATGCGGCGAATCCACTTCGATCGATCG 67
 DB 11 GCCTGCGATCTTTGGGGCCACCGGCTCATGCGGCGAATCCACTTCGATCGATCG 70
 OY 68 GGAAGGCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 127
 DB 71 GAAA-----TCCGACGCGCTCGAAGTGTGGCGTGCAGCGGAAATGCGATGCGAGA 124
 OY 128 GATGCGCGAATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 187
 DB 125 AGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 184
 OY 188 GCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 247
 DB 185 GCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 244
 OY 248 AGGCGATCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 307
 DB 245 ATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 304
 OY 308 CGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 367
 DB 305 GCGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 364
 OY 368 ACAAGGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 427
 DB 365 ACAAGGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 424
 OY 428 CCAGGATCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 487
 DB 425 CGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 484
 OY 488 ACAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 547
 DB 485 CGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 544
 OY 548 GCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 607
 DB 545 CGAAGGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 604

OY 608 TGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 667
 DB 605 TGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 664
 OY 668 AGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 727
 DB 665 AGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 724
 OY 728 CCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 787
 DB 725 CCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 784
 OY 788 CCAGGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 847
 DB 785 CCAGGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 844
 OY 848 TCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 907
 DB 845 TCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 904
 OY 908 GCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 967
 DB 905 GCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 964
 OY 968 CTTTCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1027
 DB 965 CTTTCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1024
 OY 1028 TGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1083
 DB 1025 TGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080

RESULT 3
 AAL40183
 ID AAL40183 standard; DNA; 2017 BP.
 XX
 AC AAL40183;
 XX
 DT 13-SEP-2002 (first entry)
 XX
 DE Isoprenoid related nucleic acid sequence SEQ ID No 95.
 XX
 KW Isoprenoid; Coq(10); 1-deoxyxylulose-5-phosphate synthase; Dxs;
 KW decaprenyl diphosphate synthase; ds.
 XX
 OS Sphingomonas trueperi.
 XX
 PN WO200226933-A2.
 PD 04-APR-2002.
 XX
 PF 28-SEP-2001; 2001WO-US30328.
 XX
 PR 29-SEP-2000; 2000US-236580P.
 XX
 PA (CRGI) CARGILL INC.
 XX
 PI Gokarn R, Jessen H, Zidwick MJ;
 XX
 DR MPI: 2002-416480/44.
 XX
 PT Substantially pure polypeptides having e.g.,
 PT 1-deoxyxylulose-5-phosphate synthase activity, useful for the
 PT production of isoprenoids, especially Coq(10)
 XX
 PS Claim 40; Fig 28; 246pp; English.
 XX
 CC The invention relates to methods and materials for the production of
 CC isoprenoids. More particularly the invention provides isolated nucleic
 CC acids, substantially pure polypeptides, host cells, and methods for
 CC producing various isoprenoid compounds. The polypeptides are useful for
 CC the production of isoprenoids, especially Coq(10). Expressing the pure

CC polypeptides, which has 1-deoxyxylulose-5-phosphate synthase (Dxs)
 CC activity or decaprenyl diphosphate synthase (Dds) activity, is useful for
 CC increasing production of Coq(10) in a cell having endogenous Dds
 CC activity. This polynucleotide represents a nucleic acid sequence relating
 CC to the isoprenoid production of the invention.

XX Sequence 2017 BP; 278 A; 645 C; 722 G; 372 T; 0 other;

Query Match 32.2%; Score 380.4; DB 24; Length 2017;

Best Local Similarity 60.3%; Pred. No. 1e-54; Mismatches 421; Indels 6; Gaps 1;

Matches 649; Conservative 0; Mismatches 421; Indels 6; Gaps 1;

QY 8 GCGTGTGATCTTTGGGCGCCACCGGCTCCATCGCGCAATCCACTTCGATCGTATGC 67
 DB 585 GCGTGTGATCTTTGGGCGCCACCGGCTCCATCGCGCAATCCACTTCGATCGTATGC 644
 QY 68 GGAAGGG 127
 DB 645 GAAA-----TCCGACGCGCTTCGAAAGTGTGGCGCTACCGCAAAATTGCGATGCGAGA 698
 QY 128 GACTGGCCGAATGCGCGTGGCGTGAAGCGGAGCTTGCCTACCGCGCATGAGAGACT 187
 DB 699 AGCTGGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 758
 QY 188 GCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 247
 DB 759 GCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 818
 QY 248 AGGCGATCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 307
 DB 819 ATTGGGTGTGAGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 878
 QY 308 CCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 367
 DB 879 GCGGAGGGCTCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 938
 QY 368 ACAAGGAAGCGCTGTGACGCGGAGGCGCACTTCGTATCGGACGCGGCCAGAGAAAGCG 427
 DB 939 ACAAGGAAGCGCTGTGACGCGGAGGCGCACTTCGTATCGGACGCGGCCAGAGAAAG 998
 QY 428 CCAGCATCTGCGGCTGTGACGCGGAGGCGCACTTCGTATCGGACGCGGCCAGAGAAAG 487
 DB 999 CGAGCGCTGTGCGGCTGTGACGCGGAGGCGCACTTCGTATCGGACGCGGCCAGAGAA 1058
 QY 488 ACAGGCGCTGTGCGGCTGTGACGCGGAGGCGCACTTCGTATCGGACGCGGCCAGAGAA 547
 DB 1059 CGCCAGGCGGCTGTGCGGCTGTGACGCGGAGGCGCACTTCGTATCGGACGCGGCCAG 1118
 QY 548 GCGTGTGAGCGATCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 607
 DB 1119 CGAAGGAAGCGATGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1178
 QY 608 TGGGCGAGCGGATCTGTGACGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 667
 DB 1179 TGGGCGAGCGGATCTGTGACGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1238
 QY 668 AGAGCGCGGAATCTTGGGCTGTGACGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGG 727
 DB 1239 AAGCGTGTGAGCGGCTGTGACGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1298
 QY 728 CCATGTGTGATGTGAGGCGGCTGTGACGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGG 787
 DB 1299 CCGTGTGTGATGTGAGGCGGCTGTGACGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGG 1358
 QY 788 CCGAGATCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 847
 DB 1359 CCGAGATCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1418
 QY 848 TCGCGCGGATGACCTGTGACGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 907
 DB 1419 GCGCGCGGCTGACCTGTGACGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1478
 QY 908 GCTTCCGCGCGGCTTTCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 967

DB 1479 GCTTCCCGGCGGCTGTGACGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1538
 QY 968 CCTTCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1027
 DB 1539 TTCTCAATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1598
 QY 1028 TGAACATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1083
 DB 1599 TTGAATGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1654

RESULT 4

AA140185
 ID AA140185 standard; DNA; 1191 BP.

AC AA140185;

DT 13-SEP-2002 (first entry)

DE Isoprenoid related nucleic acid sequence SEQ ID No 105.

KW Isoprenoid; Coq(10); 1-deoxyxylulose-5-phosphate synthase; Dxs; Dds;

KM decaprenyl diphosphate synthase; ds.

OS Pseudomonas aeruginosa.

PN W0200226933-A2.

PD 04-APR-2002.

PF 28-SEP-2001; 2001WO-US30328.

PR 29-SEP-2000; 2000US-236580P.

PA (CNGI) CARGILL INC.

PI Gokarn R, Jessen H, Zidwick MJ;

DR WPI; 2002-416480/44.

PT Substantially pure polypeptides having e.g.,

PT 1-deoxyxylulose-5-phosphate synthase activity, useful for the

PT production of isoprenoids, especially Coq(10)

PS Disclosure; Fig 31; 246pp; English.

XX The invention relates to methods and materials for the production of
 CC isoprenoids. More particularly the invention provides isolated nucleic
 CC acids, substantially pure polypeptides, host cells, and methods for
 CC producing various isoprenoid compounds. The polypeptides are useful for
 CC the production of isoprenoids, especially Coq(10). Expressing the pure
 CC polypeptides, which has 1-deoxyxylulose-5-phosphate synthase (Dxs)
 CC activity or decaprenyl diphosphate synthase (Dds) activity, is useful for
 CC increasing production of Coq(10) in a cell having endogenous Dds
 CC activity. This polynucleotide represents a nucleic acid sequence relating
 CC to the isoprenoid production of the invention.

XX Sequence 1191 BP; 176 A; 400 C; 413 G; 202 T; 0 other;

Query Match 24.2%; Score 286; DB 24; Length 1191;

Best Local Similarity 55.8%; Pred. No. 4.7e-39; Mismatches 455; Indels 18; Gaps 2;

Matches 597; Conservative 0; Mismatches 455; Indels 18; Gaps 2;

QY 20 TTGGGGCCACCGGCTTCATCGCAATCCACTTCGATCGTATGCGGAAGGCG---76
 DB 29 TCGGCGGCGGCGGCTTCATCGGCGGCTGAGACCGCTGAGCGTGTGACGCGTATCCGATC 88
 QY 77 GCGCGGAGCGCTTCGCGACCGCTGTGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 136
 DB 89 GTTACGAAGCGCTTCGCGGCTGTGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 148
 QY 137 AATGCGCGGCTGTGAGCGGAGCTTGGCGGCTACCGCGGCGGCGGCGGCGGCGGCGGCGG 196

```

Db 149 TCAGGACCGCCCGCTATGCGGTGTGTCGAGACGCGCGGGATGCTTGCCTGACG 208
Oy 197 GCCTGCGGAGGCGCTGGCGGGAGCGGACCGAGGTCCGGGGCGGCGGAGCCATG 256
Db 209 GCTGCTGCGCGCGGGGTATCCGACCGGGTGTCTGTTGCGGAGCGGGGTGTGCG 268
Oy 257 CCGAGGCGCGGACCGCGCGGCGGACTGAGCATGTCGACATGCGGCGCGCGCGGCG 316
Db 269 AAGTGGCGCGCGCGCGGAGAGTGAATGATGAGCGGCGCATCTGCGCGCGCGCG 328
Oy 317 TCGTGGCGGGAATGCGGCGCGGTGAAGACGCGCGGCAAGCTGGCGCTCGCCAAAGAAA 376
Db 329 TCGCGCTCGACCTGGCGCGGCTGAGAGCGCGGCAAGCGCTACTGCTGCGCAACAGAG 388
Oy 377 GCCTGTGAGCGGAGGCGGCAATCTGATGCGGACCGCGCGGAGAGAACGCGCGCATG 436
Db 389 CGCTGGTGAATGCGCGCGCGCTGTTCATGACGCGGTCAAGCGGACGCGCGCGGTCTCC 448
Oy 437 TCGCGGTGAGACGAGCACTCCGCGCTCTTCAGGCGCTGGCG-----G 481
Db 449 TCGCGATGACAGCGAGCAACAGCGATTTCCAGTCCGTGCGCGCATTTATGCCGATG 508
Oy 482 GCGAGGACAGCGGCTGCGGTGAGCGCGGTATCATACAGCGCGTCCGCGCGCGCTCCGCG 541
Db 509 GCGTGGAGCGGCGTGGCGCGTGGCGCGGATCTTTCAGCCGCTCCGCGCGCGCTCCGCG 568
Oy 542 ACTGAGCGCTGAGCGGCAATCCGCGCTGACCGGTGCGCGGCGGCGCGCGCATCCCACT 601
Db 569 AAGCGCGCGGTGAGCAACTCTGCTGGTGAAGCGCGGAGCGCGTGTGGCGCGCGCACT 628
Oy 602 GGTCCATGGGCGGAGGATCTCCATGACAGCGCGCTCATGTTTCACAGAGCGCGTGGAG 661
Db 629 GGTTCATGGGCGGTAAATTTCCGTGACTCCGCGAGCATGATGAACAGGCGCTGAAAC 688
Oy 662 TGATGAGAGCGCGCAATTTCTGCGGTGAGCGCGGAGCGGATGAGCGCGTGTCCATC 721
Db 689 TGATGAGAGCGGTGCTGCTGTTGAGCGCGGAGCGGAGCGGTGATGATCACC 748
Oy 722 GCGAATCCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 781
Db 749 CCGAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 808
Oy 782 GCGCGCGCGACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 841
Db 809 GGAATCCGAGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 868
Oy 842 TCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 901
Db 869 CCGCGCGTTCGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATG 928
Oy 902 AAGAAAGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 961
Db 929 ACGACGCGTTCGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATG 988
Oy 962 GCGCGCGCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1021
Db 989 CGGCGCATGCTGAATGCGCGGAAAGAGGTGCGCGCGCGCGCGCGCGCGCGCGCG 1048
Oy 1022 GGTTCCTGAGCATGCGCGCGGTGCTGAGAGAAAGCGCTCGCGCGCGCTTTCG 1071
Db 1049 GCTTCAGCGCATGCGCGGTATCATGAGAGCGTGTGAACGCGCGAGCGG 1098

```

RESULT 5
AAL40187

ID AAL40187 standard; DNA; 1161 BP.

AC AAL40187;

DT 13-SEP-2002 (first entry)

DE Isoprenoid related nucleic acid sequence SEQ ID No 107.

```

XX XX
KW Isoprenoid: CoQ(10); 1-deoxyxylulose-5-phosphate synthase; DXS; DKS;
NM decaprenyl diphosphate synthase; ds.
OS Streptomyces griseoliosporus.
PN WO200226933-A2.
PD 04-APR-2002.
XX 28-SEP-2001; 2001WO-US30328.
XX 29-SEP-2000; 2000US-236580P.
XX (CRGI ) CARGILL INC.
XX Gokarn R, Jessen H, Zidwick MJ;
XX WPI; 2002-416480/44.
DR
XX
XX Substantially pure polypeptides having e.g.,
PT 1-deoxyxylulose-5-phosphate synthase activity, useful for the
PT production of isoprenoids, especially CoQ(10)
XX
XX Disclosure; Fig 31; 246pp; English.
XX
XX The invention relates to methods and materials for the production of
XX isoprenoids. More particularly the invention provides isolated nucleic
XX acids, substantially pure polypeptides, host cells, and methods for
XX producing various isoprenoid compounds. The polypeptides are useful for
XX the production of isoprenoids, especially CoQ(10). Expressing the pure
XX polypeptides, which has 1-deoxyxylulose-5-phosphate synthase (Dxs)
XX activity or decaprenyl diphosphate synthase (Dps) activity, is useful for
XX increasing production of CoQ(10) in a cell having endogenous Dps
XX activity. This polynucleotide represents a nucleic acid sequence relating
XX to the isoprenoid production of the invention.
XX
XX Sequence 1161 BP; 151 A; 399 C; 449 G; 162 T; 0 other;
SQ

```

Query Match 20.9%; Score 247.2; DB 24; Length 1161;
Best Local Similarity 56.9%; Pred. No. 1.3e-32;
Matches 474; Conservative 0; Mismatches 353; Indels 6; Gaps 1;

```

Oy 231 GGTGCGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 290
Db 234 GGTGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 293
Oy 291 GTGCGCATGCTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 350
Db 294 GAACGCGCATCACCGGTTGATGCGGCTGGCCCGGAGCGTGGCGCGCGCGCGCG 353
Oy 351 CACGCTGCGGCTGCGCAACAGGAAGACCTGCTGACGCGGAGGCGCACTCTGATGCGGAC 410
Db 354 GGTGCTGCTGCTGCGCAACAGGAGTGTGATGCTGCGGCGGTCCGCTGTGAAGCGGT 413
Oy 411 GGGCCGAGGAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 470
Db 414 GGGCGAGCC-----CGGCGAGTCTGCGCGGTGAGACTTCGAGACGCGCGCGCTGTTCCA 467
Oy 471 GGGCGTGGCGGCGAGGACAGCGGCTGCTGAGAGCGCGCATCATCAACGCGCTCGGCGG 530
Db 468 GGGCGTGGCGGCGGCGCGCGCGCGCGGAGGTGCGCAAGTGTGTGACCGCGCGCGG 527
Oy 531 GCGCTTCGCGGACTGAGGCTTGAAGCGCATCCGCGCTTACACCGTGGCGGAGCGGCG 590
Db 528 CCGCTTCGCGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 587
Oy 591 CCATCGCAACTGTGCTGAGCGGCGAGCGGATCTTCATGACAGCGCGCTGATGTTCAACAA 650
Db 588 GCAACCGGACTGTGGCGGATGGCGCGGTGTGACGATCAACCTGGCGGACCTGTGACAA 647
Oy 651 GGGCGTGAAGCTGATGAGAGCGCGCGCAATTTCTGCGCTTGAAGCGGACCGGATCGAGCG 710

```


QY 705 CGAGCGGCTGCTCATCCGACATCCATGCTCATGATGGTGGCTTTCGACGGGG 764
 Db 182694 CGAAGTGGTCAATCCGACATCCGATGATACAGACATGGTGGCTACCGGAGGCTC 182635
 QY 765 CCTGATGGCCATCTCGGCCCCCGGACATGGCGGACGCCATCGGATTCGGGCTGAAGT 824
 Db 182634 CGTCTGGCGCAATGGGCAATCCGATATGCGAAGCCGCTTATTTGGTGGTTT 182575
 QY 825 GCCGGGTGCGGCGAGGTGCGGCTGCGCCGAGTGCACCTCGACAGATTCGAGCCTAC 884
 Db 182574 GCCGAGGCGCATGATTCGGGTGTGCGGACCTGATTCAGCATTTGTCCGGGCTAC 182515
 QY 885 CTTCAGAAAGCTTACAGAGAACCTTTCGCGCCCTGAGGCTTCGCGAGAGCTCATGCG 944
 Db 182514 CTTCAGAAAGCGCCGACATTTTGACCGCTTCCCTGCGGCTCGCTATGAGGCAATGAA 182455
 QY 945 GCGCGGCGGCTGTGCGGCGCGGCTTCACGCGGCGCAAGAGATTCGCGCTCATATT 1004
 Db 182454 CGCAGGCGGAGCGCGCGGCTTCGCTATTTGACGCGCGGCAAGAGCGCGGCTTC 182395
 QY 1005 CATGCGCGGAGCATGCGGCTTTCGACATGCGGCGGTGTCGAGAGACGCTCGCGG 1064
 Db 182394 TTGAGAGGAGCATATTAATTTACGACATGCGCAAAACGCTGCGCCACTGTTCGACA 182335
 QY 1065 CGTTGACGCGACCCCTGTTTCGAAAAAGTG 1095
 Db 182334 AGACTTTTCAGACGCGCATAGCGCATATAGGG 182304

RESULT 8
 AAL40188
 ID AAL40188 standard; DNA; 1185 BP.

AC AAL40188;
 XX 13-SEP-2002 (first entry)
 DE Isoprenoid related nucleic acid sequence SEQ. ID No 108.
 XX Isoprenoid; CoQ(10); 1-deoxyxylulose-5-phosphate synthase; DXS; Dds;
 KM decaprenyl diphosphate synthase; ds.
 OS Neisseria meningitidis.
 XX WO200226933-A2.
 PD 04-APR-2002.
 PF 28-SEP-2001; 2001WO-0530328.
 PR 29-SEP-2000; 2000US-236580P.
 PA (CRGI) CARGILL INC.
 PI Gokarn R, Jessen H, Zidwick MJ;
 DR WPI; 2002-416480/44.
 PT 1-deoxyxylulose-5-phosphate synthase activity, useful for the
 production of isoprenoids, especially CoQ(10)
 PT production of isoprenoids, especially CoQ(10)
 PS Disclosure; Fig 31; 246pp; English.

The invention relates to methods and materials for the production of isoprenoids. More particularly the invention provides isolated nucleic acids, substantially pure polypeptides, host cells, and methods for producing various isoprenoid compounds. The polypeptides are useful for the production of isoprenoids, especially CoQ(10). Expressing the pure polypeptides, which has 1-deoxyxylulose-5-phosphate synthase (DXS) activity or decaprenyl diphosphate synthase (Dds) activity, is useful for increasing production of CoQ(10) in a cell having endogenous Dds activity. This polynucleotide represents a nucleic acid sequence relating

CC to the isoprenoid production of the invention.
 XX Sequence 1185 BP; 255 A; 365 C; 326 G; 239 T; 0 other;
 SQ

Query Match 19.3%; Score 228.6; DB 24; Length 1185;
 Best Local Similarity 53.3%; Pred. No. 1.5e-29;
 Matches 592; Conservative 0; Mismatches 489; Indels 30; Gaps 4;

QY 9 CCTGTCATCTTTGGGGGACCGGCTCCATCGGGAATCCACTTTCGACCTGTCATGCG 68
 Db 15 CCTGACCATATTAGGAGTAGTACCGGACAGATAGGGGAAGACGGTGGACGTTGCTCCC 73
 QY 69 GAAGGCGGGGCGGAGGCTTCCGACCGTCTCTGACCGGCGGCGCAATCCGGCG 128
 Db 74 -----GCCACCCCGAATAAATTCGCGATTCGCTGCGCAGGCGCATAGCGGTGAGAA 128
 QY 129 ACTGCGCAATATGGCGGCGGTGAGAGGCGGAGCTTGCCTGACCGGCGATGAGAGACTG 188
 Db 129 ATTGGCGGCTAATGTCAAAAGCTTCCACCCGAAATATGCGCTTGCCTGCGGATGCGAACA 188
 QY 189 CCTGCGCGCTGCGCGGAGGCGCTG-----GCCGAGCGGCGACCGAGTGGCGGCG 242
 Db 189 CGCGCGCGGCTTGAAGCGCTGTGAAACGCGGACGCGGCGAGCTCAGGTTTACACGG 248
 QY 243 GCGCGAGCG---CATGCGGAGGCGCGCGGCGCGGCGGCGGAGCTGTCGCGCAT 299
 Db 249 CGCGCAGGCAATGGGTGACGTTGCGCTTGCCTGCGGACGAAATGACGCTGTCATGCGCGCAT 308
 QY 300 CGTGGGCGCGCGGCGGCGCTGCGGGAATGCGGCGGTGAGAGACGCGGCGGCGGCGTGC 359
 Db 309 CGTGGGCGGCTGCGGCGGCTGCGGCTTCCGCGCTGCGAGCGGCGGCAAAAGCAAAACATTTA 368
 QY 360 GCTGCGCAACAAGAAAAGCTGTCGAGCGGAGGCGCAATCTCTGATGCGGAGCGGCCAGGA 419
 Db 369 TCTGCGCAACAAGAGAGAGCGTGTGTTTCCGCGCGCTGTTATGGAACCGCGCGGCG 428
 QY 420 GAAGGCGGCGGAGATCTGCGCGGTGAGAGAGAGACATCCGCGGCTTTGAGGCGGTGCG 479
 Db 429 AAACGCGCGGCGAGTGTGCGCGCGTGCAGAGCAACAAACCGCTTTTCAAGTTTCC 488
 QY 480 GGGGAGAGACACGG-----CCTGCGTGCAGGCGGCTCATCATCAGCGGCTC 524
 Db 489 GCGGATTTACACAGGTGCGCTGGAACGAACAGGATGCTTCGATTCCTGACCGCTTC 548
 QY 525 CGGCGGCGGCTTCCGCGAGCTGAGAGCTTCGAGCGGATCCGCGCGCTGACCGGAGGC 584
 Db 549 CGGCGCGCGGCTTTCGACCGCGGATTTAAACAGCTGACAGCATTTAGCGCCGACAAAGC 608
 QY 585 GCAGGCGCATCCCAACTGTGTCATGCGGCGGAGATCTCCATTCGACAGCGGCTGATGTT 644
 Db 609 GGTCAAAACACCCCAATTGGCGTATGGAGCGCAAAATCTCGTGAATTCGCGCACCATGAT 668
 QY 645 CAACAAGCGGCTCGAGCTGATCGAGCGCGGAAATTCCTCGGCTTCGAGCGGAGCGGAT 704
 Db 669 GAACAAGGTTTGAAGCTGATTTGAAGCGCATTTGGCTGTTCACTGTCCGCCGCAAAACT 728
 QY 705 CGAGCGGTGCTCATCCGCAATCCATGTCATGCGATGCTGCGCTTCGAGCGGCGG 764
 Db 729 CGAAGTGTCTATTCATCCGCAATCTGTGATACAGCATGCTGCGGCTACCGCGAGGCTC 788
 QY 765 CCTATGCGGCTATCGGCGCGGCGGAGATGCGGACGCGCATTCGATTCGCGCTGAAGT 824
 Db 789 CGTGTGGCGCACTGGGCAATTCGCGATATGCGAAGCGCATTCCTTATTTGTTGGGTTT 848
 QY 825 GCGGCGGTGCGGCGAGGTCGCGCGGATGACCTGCGACATCCAGATTTGCGAGCCTAC 884
 Db 849 GCCGAGGCGCATTCATTTGCGGTGTGCGGAGACTGATTTGACCATTTGTCGGGCTGAC 908
 QY 885 CTTCAGAAAGCTTACAGAGAACCTTTCGCGCCCTGAGGCTTCGCGGAGAGAGCTATGCG 944
 Db 909 CTTCAGAAAGCGGACATTTGACCGCTTCCCTGCTGAGAGCTCGCTATGAAAGCATGAA 968
 QY 945 GCGCGGCGGCTGTGCGGCGCGGCTTCACGCGGCGCAAGAGATTCGCGCTGCAATTT 1004

Db 969 CGCAGCGAGCGCCGCGCCGCGTATGACCCGCCCAAGCGCGCGCGCCCTT 1028
 OY 1005 CATGCCGAGCGATCGGTTCTGACATGCGCGCGGTGCGAGGAGCGCTGCGGG 1064
 Db 1029 TTTGACGACGAGATTAAATTACCGACATTCGCAAAACCGCGCCATGCTTTTCA 1088
 OY 1065 CGTTTCGACCGACCCCTGTTGCGAAAGTG 1095
 Db 1089 AGACTTTTCAGACGCGCATAGCGGACATAGG 1119
 RESULT 9
 ID ABR83261 standard; DNA; 1182 BP.
 AC ABR83261;
 XX 27-AUG-2002 (first entry)
 DE High growth methanotrophic bacterial strain DNA #32.
 XX High growth methanotrophic bacterial strain: C1 carbon substrate; gene:
 KW methane; methanol; Embden-Meyerhof carbon flux pathway; 16S rRNA; 3S;
 KW pyrophosphate dependent phosphotriokinase; nitrogen-containing compound;
 KW ammonia; nitrate; nitrite; nitrogen; pigment; oxygen; landfill;
 KW methane-containing environment; waste water treatment system; isoprenoid;
 KW nitrous oxide; terpenoid; animal feed; carotenoid; exopolysaccharide.
 XX Methylomonas 16a.
 OS WO200220728-A2.
 XX 14-MAR-2002.
 XX 28-AUG-2001; 2001WO-US26827.
 XX 01-SEP-2000; 2000US-229858P.
 XX (DUPO) DU PONT DE NEMOURS & CO E. I.
 PA Koffas M, Odom JM, Schenle A;
 PI WPI: 2002-452200/48.
 DR P-PSDB; ABG61582.
 XX New high growth methanotrophic bacterial strain, useful for producing
 PT single cell proteins, grows on a C1 carbon substrate, and comprises a
 PT functional gene encoding in Embden-Meyerhof carbon pathway
 XX Claim 29; Page 140; 157p; English.
 XX The invention relates to a high growth methanotrophic bacterial strain,
 CC which grows on a C1 carbon substrate e.g. methane and methanol, and
 CC comprises a functional Embden-Meyerhof carbon flux pathway comprising a
 CC gene coding a pyrophosphate dependent phosphotriokinase enzyme or a 16S
 CC rRNA. The bacterial strain is useful for the production of single cell
 CC protein and for the biotransformation of a nitrogen-containing compound,
 CC e.g. ammonia, nitrate, nitrite or nitrogen. It is also useful for the
 CC production of a feed product comprising a protein, carbohydrates and a
 CC pigment and for reducing oxygen demand, for removing nitrates and
 CC nitrates in methane-containing environments such as landfills, waste
 CC water treatment systems or anywhere that methane, oxygen and nitrates are
 CC present. The bacterial strain of the invention can be used as a
 CC denitrifying agent for the conversion of nitrate or nitrite to nitrous
 CC oxide with methane or methanol as a carbon source. It is also used in the
 CC production of biomass including proteins, carbohydrates and a wide
 CC variety of pigments (particularly for isoprenoid pigments for the
 CC purpose of generating animal feeds), in production of terpenoid and
 CC carotenoid compounds, useful as pigments and as monomers in polymeric
 CC materials and in production of exopolysaccharides at high levels.
 CC Sequences ABR83230-ABR83270 represent high growth methanotrophic
 CC bacterial strain DNA of the invention.

XX SQ Sequence 1182 BP; 270 A; 301 C; 347 G; 264 T; 0 other;
 Query Match 19.0%; Score 224.2; DB 24; Length 1182;
 Best Local Similarity 56.5%; Pred. No. 8,2e-29;
 Matches 446; Conservative 0; Mismatches 328; Indels 15; Gaps 1;
 OY 277 GCCGACTGAGCATTGCTGCGCATTCGTGAGCGCGGCGGCTGCGCGGATGCGGCG 336
 Db 283 GTGATACGCGTATGCGGCGTATCGTGGCGGCGCGCGGATGTTGCGGACCTTGGCGCG 342
 OY 337 CTGAAGCAGCGCGCGACGCTGCGCTCGCCAAAGAAAGAAAGCTTGAGCGGCGGCA 396
 Db 343 GCCAAGCGCGGCAAAACCGTCTGTGGCCAAAGAAAGAGCTTGATGCGGAGCA 402
 OY 397 CTCCTGATGCGGAGCGCGCGGAGAAAGCGCGGCGGATGCGCGGAGCGGAGCA 456
 Db 403 ATCTTCATGCGAGCGCGTCAAGATTCGCGGCTGTGTGCTCGGATGACAGCGAGCA 462
 OY 457 TCCGCGCTTTCAGGCGCTGCGCGGCGGCGAGCAGCGGCGCTGC-----GTC 501
 Db 463 AAGCGCATCTTCAGTGCATGCGCGGCGGCTTATGCGCCAGCGCATACAGCAAGAGCGG 522
 OY 502 GAGCGGCTATATCATCAGCGGCTCGCGGCGGCGCTTCCGCGACTGAGCGCTCGAGCGCATC 561
 Db 523 CGCCGCAATTTTATGACCGCTTCGCGTGGCCCATTTGCGAGCGCGCATAGAAAGCTTG 582
 OY 562 CGCGCGTCAAGCGGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 621
 Db 583 TCCAGCGTACCGCGGATAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 642
 OY 622 TCCATCGACAGCGCGCTCGATGTTCAAGAGCGGCGTGCAGCTGATGAGAGCGCGCATTC 681
 Db 643 TCGGTCGATTCGCGCCACCATGATGAACAAAGGCTGGAAGCTGAGCAAGCTGCTGTTG 702
 OY 682 TTCGGCTTGAGCGCGGAGCGGATCGAGCGGCTGTCATTCGCGATTCATGCTCATGCG 741
 Db 703 TTCAACATGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 762
 OY 742 ATGTGCGCTTTCGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 801
 Db 763 ATGTGCGCTTTCGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 822
 OY 802 GCCATCGATTCGCGCTGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 861
 Db 823 CCGATGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 882
 OY 862 CTCGACAGATTCGAGCGCTTCACCTTCGAAAGCGCTGACGAGAGAGCTTTCGCGCGT 921
 Db 883 ATTTTCGAAGTGGGCGACATGATTTGAAACCGGACTTGAAGAGGTTCTGCTGCG 942
 OY 922 AGCTTTCGCGAGAGCTGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 981
 Db 943 AAGATGCGCTTATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1002
 OY 982 AAGAGATTCGCGCTGATTCATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1041
 Db 1003 AATGAATTCGCTGCGAGCGGCTTTTAATGAAGATCAATTCATGACATCGCGGCTC 1062
 OY 1042 GTGCTCGAG 1050
 Db 1063 ATCATCGAG 1071
 RESULT 10
 ID AAD35500 standard; DNA; 1182 BP.
 AC AAD35500;
 XX 25-JUL-2002 (first entry)
 DE Methylomonas 16a sp. D-1-deoxyxylose-5-phosphate reductoisomerase gene.

PF 29-AUG-2001; 2001WO-US26852.
 XX
 PR 01-SEP-2000; 2000US-229907P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.
 XX
 PI Cheng Q, Kofas M, Norton KC, Odom JM, Picataggio SK, Rouviere PE;
 PI Schenzle A, Tomb J;
 XX
 DR WPI: 2002-383051/41.
 DR P-PSDB; AA080326.
 XX
 PT Novel nucleic acid molecule encoding a isoprenoid biosynthetic enzyme,
 PT isolated from *Methylobionas 16a*, useful for the production of isoprenoid
 PT compounds -
 XX
 PS Claim 2; Page 68; 84pp; English.
 XX
 CC The present invention relates to a new nucleic acid molecule encoding
 CC an isoprenoid biosynthetic enzyme isolated from *Methylobionas 16a*.
 CC The invention is useful for obtaining a nucleic acid molecule
 CC encoding an isoprenoid compound biosynthetic enzyme, and for the
 CC microbial production of isoprenoid compounds. The molecules of the
 CC invention are also useful for regulating isoprenoid biosynthesis in an
 CC organism and for producing recombinant organisms for producing various
 CC isoprenoid compounds. The nucleic acid is also useful for feed additive,
 CC for the production of keratensoids and their derivatives, isoprenoid
 CC intermediates, and as pure products useful as pigments, flavours and
 CC fragrances. The present nucleic acid sequence encodes the *Methylobionas*
 CC 16a open reading frame 2 (ORF2) dxr (1-deoxyxylulose-5-phosphate
 CC reductoisomerase enzyme) protein of the invention, as described above.
 XX
 SQ Sequence 1182 BP; 270 A; 301 C; 347 G; 264 T; 0 other;

Query Match 19.0%; Score 224.2; DB 24; Length 1182;
 Best Local Similarity 56.5%; Pred. No. 8.2e-29;
 Matches 446; Conservative 0; Mismatches 328; Indels 15; Gaps 1;

QY 277 GCCGACTGACACATGTCGGCCATGTCGGCGCGCGCGCTCGTCCCGGATCGCGCG 336
 DB 283 GTGATACGCGTATGCGGCTATCGTGGCGCGCGGATGTTGTCGACCTTGGCGCG 342
 QY 337 CTGAGCAGCGCGCGCGCGCTCGCGCCACAGAGAAAGCTTCTGACGCGAGCGCAA 396
 DB 343 GCCAAGCGCGCGCGCGCGCTGTCGCGCAACAGAGAAAGCTTCTGATGTCGCGCA 402
 QY 397 CTCCTATGCG 456
 DB 403 ATCTTATGCG 462
 QY 457 TCCGCGCTGTCAGCG 501
 DB 463 AACGCGATCTTTCATGTCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 522
 QY 502 GAGCGCGTCATCATCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 561
 DB 523 CGCGCGATTTTATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 582
 QY 562 CGCGCGTCGACCG 621
 DB 583 TCCAGCGTCAGCG 642
 QY 622 TCCATGACAGCG 681
 DB 643 TCGGTGATTCG 702
 QY 682 TTGCGCTGAGCG 741
 DB 703 TTCAACATGAGCG 762
 QY 742 ATGCTGGGCTTTCG 801
 DB 763 ATGCTGGGCTTTCG 822

QY 802 GCCATCGGATTCG 861
 DB 823 CCATATGCG 882
 QY 862 CTCGACAGATTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 921
 DB 883 ATTTGCAAGTATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 942
 QY 922 AGGCTTCG 981
 DB 943 AGATTCGCTTATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1002
 QY 982 AAGGATTCG 1041
 DB 1003 AATGAATTCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1062
 QY 1042 GTGCTCGAG 1050
 DB 1063 ATCATCGAG 1071

RESULT 12
 ID ABL53861 standard; DNA; 11575 BP.
 XX
 AC ABL53861;
 XX
 DT 24-JUN-2002 (first entry)
 XX
 DE Plasmid pTUS75:dxs:dxr:Tn5kn.
 XX
 KW Plasmid pTUS75:dxs:dxr:Tn5kn; vector; limonene synthase;
 XX monoterpenes; cyclic terpenoids; ds.
 OS Chimeric - *Methylobionas* sp.
 OS Chimeric - Unidentified.
 XX
 PN WO200220815-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 28-AUG-2001; 2001WO-US26853.
 XX
 PR 01-SEP-2000; 2000US-229858P.
 PR 01-SEP-2000; 2000US-229907P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.
 XX
 PI Dicosimo DJ, Kofas M, Wang S;
 PI WPI: 2002-339805/37.
 XX
 PT Producing cyclic terpenoids, e.g. monoterpenes, which are useful in the
 PT fragrance or pharmaceutical industry, by employing a transformed
 PT methanotrophic bacterium that metabolize single carbon substrates as a
 PT sole carbon source -
 XX
 PS Example 6; Page 53-60; 63pp; English.
 XX
 CC The present sequence is that of vector plasmid pTUS75:dxs:dxr:Tn5kn.
 CC This is a derivative of RK2, a broad host-range plasmid, modified to
 CC include kanamycin resistance and 2 *Methylobionas* genes: dxs encoding
 CC 5'-deoxy-D-xylulose synthase; and dxr encoding 5'-deoxy-D-xylulose
 CC reductoisomerase. A *Mentha spicata* limonene synthase gene (see
 CC ABL53856) encoding a truncated enzyme (see ABL5749) was ligated
 CC into the vector, creating plasmid pDH3, in which the limonene
 CC synthase gene was flanked on the 5' end by dxs and on the 3' end by
 CC dxr. pDH3 was transferred to *Methylobionas* sp. 16a by triparental
 CC conjugal mating. Cells that received the limonene synthase gene
 CC demonstrated to produce limonene. This is an example of the method
 CC of the invention, in which a C1 metabolizing host cell is
 CC transformed with a gene encoding a cyclic terpene synthase, and is

CC used to produce a monoterpane.

XX Sequence 11575 BP; 2395 A; 3273 C; 3548 G; 2359 T; 0 other:

Query Match 19.0%; Score 224.2; DB 24; Length 11575;

Best Local Similarity 56.5%; Pred. No. 6,9e-29;

Matches 446; Conservative 0; Mismatches 328; Indels 15; Gaps 1;

```

QY 277 GCGGACATGAGACATGTCGGCCATGTCGGGCGCCGCGGCGCTGCGGAGCAATGCGGCGG 336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3339 GTGATACGATGATGCGGCGCTATGTCGCGCGCGCGCGATTTGTTCCGACCTTGCGCG 3398
QY 337 CTGAAGACGCGCGCACGCTGCGCGCTGCCAACAAGAAAGCTCTGACGCGAGCGCA 396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3399 GCCAAGCGCGGCAAAACCGCTGCTGTGGCCAAAGAAAGCTTGATGATCGGGACAA 3458
QY 397 CTCTGTATGCGGAGCGCGCGCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3459 ATCTTCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3518
QY 457 TCGCGGCTCTTTCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 501
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3519 AAGCCCATCTTTCAGTGCATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3578
QY 502 GAGCGCGCTCATCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3579 CGCGCGCATTTTATTTAGACGCGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3638
QY 562 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 621
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3639 TCCAGCGCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3698
QY 622 TCCATCAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 681
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3699 TCGGTCGATTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3758
QY 682 TTGCGGCTTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 741
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3759 TTCAACATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3818
QY 742 ATGTGTGCGCTCTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 801
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3819 ATGTGTGCGCTCTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3878
QY 802 GCCATCGGATTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 861
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3879 CCGATACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3938
QY 862 CTGCGCAGATGTCGAGCGCTTCCAGAGCGCTGACGAGGAAAGCTTTCCGCGCGCTG 921
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3939 ATTTTCGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3998
QY 922 AGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 981
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3999 AGATTGCGCTTATGAGCGCATCACTGTGTGATATGCAAGGATTTAGACGCGCG 4058
QY 982 AAGGAGATCGCGCTGATCATTTATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1041
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4059 AATGAATTTGCTGTGCGAAGCGTTTAAATGAAGAAAGTCAATTCAGTACATCGCGCG 4118
QY 1042 GTGCTCGAG 1050
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4119 ATCATCGAG 4127
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 13

AA253883 standard; DNA; 888 BP.

XX AA253883;

DT 21-MAR-2000 (first entry)

DE Neisseria meningitidis ORF 572 partial DNA sequence SEQ ID NO:1715.

XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;

KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;

KM antibacterial; gene therapy; ds.

XX Neisseria meningitidis.

XX WO957280-A2.

XX 11-NOV-1999.

XX 30-APR-1999; 99WO-US09346.

XX 01-MAY-1998; 98US-0083758.

XX 31-JUL-1998; 98US-0094869.

XX 02-SEP-1998; 98US-0098994.

XX 02-SEP-1998; 98US-0099062.

XX 09-OCT-1998; 98US-0103749.

XX 09-OCT-1998; 98US-0103794.

XX 09-OCT-1998; 98US-0103796.

XX 25-FEB-1999; 99US-0121528.

XX (CHIR) CHIRON CORP.

XX (GENO-) INST GENOMIC RES.

XX Fraser C, Galeotti C, Grandi G, Hickey E, Masiagnani V, Mora M;

XX Petersen J, Piza M, Rappunli R, Rattl G, Scalato E, Scarselli M;

XX Tettelin H, Venter JC;

XX WPI; 2000-062150/05.

XX P-PSDB; AAY75121.

XX Novel Neisserial polypeptides predicted to be useful antigens for

XX vaccines and diagnostics

XX Claim 7; Page 879; 1453pp; English.

XX AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941

XX represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides

XX and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent

XX PCR primers used in the exemplification of the present invention. The

XX polypeptides, the polynucleotides, antibodies and compositions of

XX the invention can be used as vaccines, as diagnostic reagents, and as

XX immunogenic compositions. The polypeptides can be used in the

XX manufacture of medicaments for treating or preventing infection due to

XX Neisserial bacteria (e.g. meningitis and septicemia), to detect the

XX presence of Neisseria bacteria, or to raise antibodies. They may also

XX be used to screen for agonists or antagonists, which may themselves

XX have use as antibacterial agents. The polynucleotides of the invention

XX may also be used in gene therapy protocols.

XX Sequence 888 BP; 186 A; 280 C; 242 G; 180 T; 0 other:

Query Match 18.4%; Score 217.2; DB 21; Length 888;

Best Local Similarity 55.2%; Pred. No. 1.2e-27;

Matches 454; Conservative 0; Mismatches 353; Indels 15; Gaps 1;

```

QY 289 ATGTGCGCATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 ATGTGCGCATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
QY 349 CGCAGCGTGGCGCTGCGCGCAAGAGAAAGCTGTGACGCGGCGCAACTCTGATGCGG 408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 AAACCATTTTATGTGCGCAAAACAAGAAACGCTGTGTTTCCGCGCGCTTGTATGGA 120
QY 409 ACGGCGCGAGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 ACCGCGCGTGCAGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
QY 469 CAGCGCGTGGCGCGCGAGAGACAGG-----CCTGCGTCGACGCGCGTCATC 513
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 CAACTTTTGGCGCGCATTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240

```

QY 514 ATCAGCGGCTCCGCGCGGCGCTTCCGACGTGAGGCTCGAGGCGATCCGCGCTGCACC 573
 DB 241 CTGACCGGCTTCCGCGCGGCGCTTCTGACGCGGATTTAAACAGCTTCAGCCGATTAAG 300
 QY 574 GTGGCCGAGCGGAGGCGGATCCCAACGTGTCATGGCCGAGATTCATCGACAGC 633
 DB 301 CCGCGCCAAAGCGGCTCAAAACCCCAATTCGCGTGGGCGCAAAATCTCCGCGATTC 360
 QY 634 GCGTCGATGTCACAAAGCGGCTCGAGCTGATGAGAGCGGCGCAATTCCTCGGCTCGAG 693
 DB 361 GCCACCATATGATACAAAGTTGGAGCTGATTTGAAGCGCATTTGGCTTTCAACTGCC 420
 QY 694 CCGGACCGGATGAGCGGCTGCTGCTCCGCAATCCATGCTCCATGCTGATGGGCTTC 753
 DB 421 CCGGACAACTGCAAGTCGATCATCCGCAATCCGATATACAGCATGGGCGCTAC 480
 QY 754 TGGAGCGGGGCTGATGCGCCATCTCGGCGCCCGGACATGGCGGCAACGCGCATGATTC 813
 DB 481 CCGGAGGCTCGCTGCTGCGCAACGCGCAATCCGATATGCGAAGCGCCATCGCTTAT 540
 QY 814 GCGCTGACCTGCGCGGCTGCGGCGGAGTGGCCGCGCGATGCGACCTCGCACAGATT 873
 DB 541 TGTTCGGCTTTCGCGAGGCGCATGCTGCGGTGCGGCGACCTGATTTGAGCGATTG 600
 QY 874 GCGAGCTCAGCTTCGAGAGCGCTGACGAGAGCGCTTCGCGCGCTGAGGCTTCGCGCA 933
 DB 601 TCGCGCTGACCTTCCAAAAGCGCGACTTTGACGCGCTTCGCGTCCGCTGAGGCGTGCCTAT 660
 QY 934 GACGTATGCGGCGGCGGCGCTGTGCGGCGCGCTTCACAGCGGCGCAAGAGATCGG 993
 DB 661 GAAGCATATACGCGAGCGGAGCGCGCTGCTGATTAACGCGCGCAAGAGCGCGC 720
 QY 994 CTGATCATTTATCGCGCGAGCATCGGCTTCGAGCATGGCGGCGGCTGCGAGAG 1053
 DB 721 GTGCGCGCTTTTGGACGAGCATTAAGTTACGACATTCGCAAAACGTCGCCAC 780
 QY 1054 ACCTGCGGCGGCTTTCGACCGAGCGCGCTGCTGCGAAAGTG 1095
 DB 781 TGTCTTGCAAGACTTTTCAGAGCGCATAGGCGGATATAGG 822
 RESULT 14
 AAH52058
 ID AAH52058 standard; DNA; 1311 BP.
 AC AAH52058;
 XX 04-SEP-2001 (first entry)
 DT Mycobacterium tuberculosis potential drug target gene SEQ ID 112.
 XX Mycobacterium tuberculosis potential drug target gene SEQ ID 112.
 DE Drug target: growth; organism viability; characterisation; ds.
 KW Mycobacterium tuberculosis.
 OS Mycobacterium tuberculosis.
 XX WO200135317-A1.
 PN 17-MAY-2001.
 XX 13-NOV-2000; 2000WO-US31152.
 PF 12-NOV-1999; 99US-0165086.
 PR 12-NOV-1999; 99US-0165124.
 PR 01-FEB-2000; 2000US-0179531.
 XX (REGC) UNIV CALIFORNIA.
 PA Eisenberg D, Rotstein SH, Marcotte EM;
 PI MPI: 2001-329193/34.
 DR P-PSDB; MAG81207.
 XX

PT Identifying nucleotide or polypeptide sequence for use as drug target,
 PT involves providing algorithm that analyzes a functional relationship
 PT between nucleotide or polypeptide sequences, and comparing the
 PT sequences
 PS Disclosure; Page 125-126; 207pp; English.
 XX
 CC This invention relates to a method for identifying a nucleotide or
 CC polypeptide sequence that may be a drug target, or essential for growth
 CC or viability of an organism. Polynucleotide sequences AAH51947 - AAH52092
 CC represent DNA encoding proteins MAG81096 - MAG81241, Mycobacterium
 CC tuberculosis proteins which are potential drug targets. The DNA and
 CC protein sequences are used to illustrate the method of the invention. The
 CC method involves providing an unknown nucleotide or polypeptide sequences,
 CC and comparing it to a number of sequences along with at least one
 CC algorithm capable of analysing a functional relationship between
 CC nucleotide and polypeptide sequences. The method is useful for
 CC characterising the function of nucleic acids and polypeptides that may be
 CC useful as a target for a drug or essential for the growth or viability of
 CC an organism.
 CC
 SO Sequence 1311 BP; 201 A; 429 C; 461 G; 220 T; 0 other;
 Query Match 18.2%; Score 215.6; DB 22; Length 1311;
 Best Local Similarity 54.2%; Pred. No. 2.2e-27;
 Matches 462; Conservative 0; Mismatches 384; Indels 6; Gaps 1;
 QY 201 GCGCGAGGCGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 260
 DB 282 GCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 341
 QY 261 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 320
 DB 342 GGTGAG 401
 QY 321 GCCCGAATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 380
 DB 402 ACCGAGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 461
 QY 381 CGTGAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 440
 DB 462 GGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 515
 QY 441 GGTGAG 500
 DB 516 GGTGAG 575
 QY 501 CGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 560
 DB 576 CGCGAAGCTGCTGCTACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 635
 QY 561 CCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 620
 DB 636 CGAGCATGTCACCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGGA 695
 QY 621 CTCGATGAG 680
 DB 696 CAGGCTGAGATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 755
 QY 681 CTTGCGCTTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 740
 DB 756 GTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 815
 QY 741 GATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 800
 DB 816 GATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 875
 QY 801 CCGCATGAGATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 860
 DB 876 ACCGATTTGCTAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 935
 QY 861 CTTGCGGAGATTCGAGCGGCTGCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 920
 XX

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2003, 00:00:39 ; Search time 3441 Seconds

(without alignments)
9996.953 Million cell updates/sec

Title: US-09-673-198-31

Perfect score: 1182

Sequence: 1 atgcgcagccgtcgtacctt.....gtctccgcagcagaagaag 1182

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_om:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1182	100.0	1182	6 E59203	E59203 Method for
2	1182	100.0	1182	6 E64384	E64384 Process of
3	480.6	40.7	333800	1 SME591792	AL591792 Sinorhizo
4	434.2	36.7	9436	1 AE008174	AE008174 Agrobacte
5	434.2	36.7	10301	1 AE009208	AE009208 Agrobacte
6	352.6	29.8	12977	1 AE005865	AE005865 Caulobact
7	305.4	25.8	11150	1 AE012791	AE012791 Chlorobiu
8	292.2	24.7	201050	1 AL646064	AL646064 Ralstonia
9	290.6	24.6	10029	1 AE012236	AE012236 Xanthomon
10	286	24.2	14537	1 AF282879	AF282879 Pseudomon
11	286	24.2	14537	1 AE004785	AE004785 Streptomy
12	278.6	23.6	39744	1 SC5H4	AL355913 Streptomy
13	255.2	21.6	10242	1 AE011773	AE011773 Xanthomon
14	247.2	20.9	1161	1 AB049187	AB049187 Streptomy
15	238.2	20.2	13538	1 AE002375	AE002375 Neisseria
16	238.2	20.2	349980	6 AX043922	AX043922 Sequence
17	228.6	19.3	340806	1 NMA122491	AL162752 Neisseria
18	224.2	19.0	1182	6 AX394964	AX394964 Sequence
19	224.2	19.0	1182	6 AX398486	AX398486 Sequence
20	224.2	19.0	1182	6 AX398794	AX398794 Sequence
21	224.2	19.0	11575	6 AX398630	AX398630 Sequence
22	215.6	18.2	39991	1 MTCY274	Z74024 Mycobacteri
23	214.6	18.2	14715	1 AE007117	AE007117 Mycobacte
24	211	17.9	299950	1 AP005372	AP005372 Thermosyn
25	201.2	17.0	34083	1 AF124757	AF124757 Zymomonas
26	196.4	16.6	1167	1 ZMO250714	AJ250714 Zymomonas
27	195.4	16.5	10487	1 AE003942	AE003942 Xylella
28	183	15.5	11576	1 AE001994	AE001994 Delnoccoc
29	181	15.3	318200	1 MEPRN6	AL583922 Mycobacte
30	179.2	15.2	1115	6 AX432668	AX432668 Sequence
31	176.2	14.9	36985	1 MSGB1529CS	L76824 Mycobacteri
32	172.4	14.6	5476	1 SLE250721	AJ250721 Synechoco
33	166	14.0	208780	1 BSUB0009	Z99112 Bacillus su
34	152.6	12.9	146271	1 SYCSLRB	D64000 Synechocyst
35	151.8	12.8	303249	1 AP001515	AP001515 Bacillus
36	146.6	12.4	250050	1 AL591978	AL591978 Listeria
37	145.8	12.3	1176	6 AX122293	AX122293 Sequence
38	145.8	12.3	337200	1 AP005280	AP005280 Coryneb
39	145.8	12.3	349980	6 AX127149	AX127149 Sequence
40	145.8	12.3	349980	6 AX127150	AX127150 Sequence
41	142.8	12.1	22997	1 AE008705	AE008705 Salmonell
42	141.2	11.9	268050	1 AL627266	AL627266 Salmonell
43	139.6	11.8	10808	1 AE001235	AE001235 Treponema
44	139.6	11.8	348550	1 AP003596	AP003596 Nostoc sp
45	132.8	11.2	1706	8 AF367205	AF367205 Oryza sat

ALIGNMENTS

RESULT 1
LOCUS E59203 1182 bp DNA linear PAT 31-JAN-2002
DEFINITION Method for searching antibacterial or herbicidally active compound.
ACCESSION E59203
VERSION E59203.1 GI:18622474
KEYWORDS JP 2000300257-A/23.
SOURCE Rhodobacter sphaeroides.
ORGANISM Rhodobacter sphaeroides
Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
Rhodobacter.
REFERENCE 1 (bases 1 to 1182)
Miyake,K., Hashimoto,S., Motoyama,H., Ozaki,A., Set,H., Kuziyama,T.
and Takahashi,S.
TITLE Method for searching antibacterial or herbicidally active compound

JOURNAL									
Patent: JP 2000300257-A 23 31-OCT-2000;									
SANKYO CO. LTD									
OS Rhodobacter sphaeroides									
PN 31-OCT-2000									
PD 31-OCT-2000									
PF 12-APR-1999 JP 1999104590									
PR									
PI KOICHIRO MIYAKE,SHINICHI HASHIMOTO, HIROAKI MOTOMAYA, AKIO									
OZAKI,									
PI HARUO SEFO,									
PI TOMOHISA KUZUYAMA, SHUNJI TAKAHASHI									
PC C12N15/09, A01N57/12, C12N1/21, C12N9/00, C12P23/00, C1201/18// PC									
(C12N1/21, C12R1:18), (C12P23/00, C12R1:19), (C12P23/00, C12R1:18), PC									
C12N15/00									
CC									
FH Key									
FT CDS Location/Qualifiers									
1. 1182									
/organism="Rhodobacter sphaeroides"									
/db_xref="taxon:1063"									
BASE COUNT 185 a 408 c 417 g 172 t									
ORIGIN									
Query Match 100.0%; Score 1182; DB 6; Length 1182;									
Best Local Similarity 100.0%; Pred. No. 1,1e-123;									
Matches 1182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	ATGCCAGCCTTCATCTTTGGGACCAACGGCTCCATCGGGAATCCACCTTCGACCTC	60						
Db	1	ATGGCAGCCTTCATCTTTGGGACCAACGGCTCCATCGGGAATCCACCTTCGACCTC	60						
QY	61	GTCATCGGAAGGGCGGGCGGCGAGCGCTTCGCACCGCTGCTGTACCGGGCGCGCAAC	120						
Db	61	GTCATCGGAAGGGCGGGCGGCGGCGAGCGCTTCGCACCGCTGCTGTACCGGGCGCGCAAC	120						
QY	121	ATCCGGCAGCTGGCGGAATGCGGGTGGCGTGAAGGGGAGCTTGGCTACCGCGCAT	180						
Db	121	ATCCGGCAGCTGGCGGAATGCGGGTGGCGTGAAGGGGAGCTTGGCTACCGCGCAT	180						
QY	181	GAGGACTGCTGCCCGCGCTGCGCAGAGCGCTGGCCGGGACCGGACCCAGGTCGGGGCG	240						
Db	181	GAGGACTGCTGCCCGCGCTGCGCAGAGCGCTGGCCGGGACCGGACCGGAGTCGGGGCG	240						
QY	241	GGGGCGGAGCCATGCGCGGAGGGCCGCCACCGCGCGCGACTGAGACATGTGCGCATC	300						
Db	241	GGGGCGGAGCCATGCGCGGAGGGCCGCCACCGCGCGCGACTGAGACATGTGCGCATC	300						
QY	301	GTCGCGCGCGCGGCTGTGTCGCCGAATGCGGGCGCTTAACACAGCGCGCAGCTGGCG	360						
Db	301	GTCGCGCGCGCGGCTGTGTCGCCGAATGCGGGCGCTTAACACAGCGCGCAGCTGGCG	360						
QY	361	CTCGCCCAACAGGAAGCTCTGTGACGGCAGGGCACTCTGTATGGGAGCGGCCACAGAG	420						
Db	361	CTCGCCCAACAGGAAGCTCTGTGACGGCAGGGCACTCTGTATGGGAGCGGCCACAGAG	420						
QY	421	AACGGCGCAGATCTGCGCGGTGACAGCGAGCACTCGCGGCTTTTAAAGCGCTGGGG	480						
Db	421	AACGGCGCAGATCTGCGCGGTGACAGCGAGCACTCGCGGCTTTTAAAGCGCTGGGG	480						
QY	481	GGCGAGGACGCGCTGTGTGAGCGCGCTCATCATCAGGCGCTCGGGCGGGCGTTCCGC	540						
Db	481	GGCGAGGACGCGCTGTGTGAGCGCGCTCATCATCAGGCGCTCGGGCGGGCGTTCCGC	540						
QY	541	GACTGGAGCCTTGAGCGCATCGCGGCTTGACACCGTGGCCGAGGCGCAGCCATCCCAAC	600						
Db	541	GACTGGAGCCTTGAGCGCATCGCGGCTTGACACCGTGGCCGAGGCGCAGCCATCCCAAC	600						
QY	601	TGCTCATGGGGCAGCGGATCTCATTCGACAGCGCGCTCGATCTTAAACAAGCGCTCGAG	660						
Db	601	TGCTCATGGGGCAGCGGATCTCATTCGACAGCGCGCTCGATCTTAAACAAGCGCTCGAG	660						
QY	661	CTGATCGAGCGCGCAATTTCTTGCGCTTGAAGCGGACCGGATGAGGCGGCTGTCAT	720						

[illegible]

/organism="Rhodobacter sphaeroides"
/db_xref="taxon:1063"
BASE COUNT 185 a 408 c 417 g 172 t
ORIGIN

Query Match 100.0%; Score 1182; DB 6; Length 1182;
Best Local Similarity 100.0%; Pred. No. 1,1e-123;
Matches 1182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGAGGCTGTCATCTTTGGGGCCACCGGCTCCATGGCGAATCCACCTTCGACCTC 60
DB 1 ATGCGAGGCTGTCATCTTTGGGGCCACCGGCTCCATGGCGAATCCACCTTCGACCTC 60
QY 61 GTCATCGGAGAGGG 120
DB 61 GTCATCGGAGAGGG 120
QY 121 ATCCGGGAGCTGGCCGAATATGGCGCTGGCGTGAAGGGGAGCTTCGCTACCGGCGCAT 180
DB 121 ATCCGGGAGCTGGCCGAATATGGCGCTGGCGTGAAGGGGAGCTTCGCTACCGGCGCAT 180
QY 181 GAGGAGTGGCTGCTGGCGGCTGGCGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 240
DB 181 GAGGAGTGGCTGCTGGCGGCTGGCGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 240
QY 241 GGGGGCAGAGCCATCGCCAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 300
DB 241 GGGGGCAGAGCCATCGCCAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 300
QY 301 GTGGGGCGCGGGGGCTGTGGTGGGGGAATGCGGGGCTTCAAGCAGCGCGCGCATGGGG 360
DB 301 GTGGGGCGCGGGGGCTGTGGTGGGGGAATGCGGGGCTTCAAGCAGCGCGCGCATGGGG 360
QY 361 CTCGGCAACAGGAAGGCTCTGAGGAGGAGGAGCACTCTGATGGGAGCGGGCGGAGAG 420
DB 361 CTCGGCAACAGGAAGGCTCTGAGGAGGAGGAGCACTCTGATGGGAGCGGGCGGAGAG 420
QY 421 AACGGGCGCAGATCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
DB 421 AACGGGCGCAGATCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
QY 481 GGGGAGGAGACGGCTGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
DB 481 GGGGAGGAGACGGCTGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
QY 541 GACTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
DB 541 GACTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
QY 601 TGGTCCATGGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
DB 601 TGGTCCATGGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
QY 661 CTGATGAGAGCGGGCAATTTCTTGGCTTGGAGCCGAGCCGAGTTCAGAGGCGCTCGAC 720
DB 661 CTGATGAGAGCGGGCAATTTCTTGGCTTGGAGCCGAGCCGAGTTCAGAGGCGCTCGAC 720
QY 721 CCGCAATCATGCTGTCATGAGATGATGGGGTTCGAGAGGGGGGCGTGAATGCGCATTC 780
DB 721 CCGCAATCATGCTGTCATGAGATGATGGGGTTCGAGAGGGGGGCGTGAATGCGCATTC 780
QY 781 GGGGGCGCGGAGATGGCGCAGCGCATGGAATTCGAGGAGGAGGAGGAGGAGGAGGAG 840
DB 781 GGGGGCGCGGAGATGGCGCAGCGCATGGAATTCGAGGAGGAGGAGGAGGAGGAGGAG 840
QY 841 GTGGCGGCTGGCGGAGTGAATTCGAGCAGATTCGAGGAGGAGGAGGAGGAGGAGGAG 900
DB 841 GTGGCGGCTGGCGGAGTGAATTCGAGCAGATTCGAGGAGGAGGAGGAGGAGGAGGAG 900
QY 901 GAGGAGCGCTTTCGGCGGCTTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
DB 901 GAGGAGCGCTTTCGGCGGCTTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960

QY 961 GGGCGCGCTTCAACGCGGCGCAGAGATGCGCTGATCATTTTCATCGCCGAGCATC 1020
DB 961 GGGCGCGCTTCAACGCGGCGCAGAGATGCGCTGATCATTTTCATCGCCGAGCATC 1020
QY 1021 TGGTTTCTGACATGGCGGGGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 1021 TGGTTTCTGACATGGCGGGGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1081 CTGTTGGGAAGATGGCGGCGCTTGGAGAGAGTCTGCGCATTCGATCGCTCGG 1140
DB 1081 CTGTTGGGAAGATGGCGGCGCTTGGAGAGAGTCTGCGCATTCGATCGCTCGG 1140
QY 1141 AGAGCGCAG 1182
DB 1141 AGAGCGCAG 1182

RESULT 3 SME591792 LOCUS

DEFINITION Sinorhizobium meliloti 1021 complete chromosome; segment 11/12.
ACCESSION AL591792 AL591688
VERSION AL591792.1 GI:15075850
KEYWORDS
SOURCE
ORGANISM Sinorhizobium meliloti.
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
Rhizobiaceae; Sinorhizobium.

REFERENCE

AUTHORS Capela, D., Barloy-Hubler, F., Gouzy, J., Bothe, G., Ampe, F., Batut, J., Boistard, P., Becker, A., Boutry, M., Cadieu, E., Dreno, S., Gloux, S., Godrie, T., Goffeau, A., Kahn, D., Kiss, E., Lelaune, V., Masny, D., Pohl, T., Portetelle, D., Puehler, A., Purnelle, B., Ramsperger, U., Renard, C., Rhébaud, P., Vandenbol, M., Weidner, S. and Galibert, F.
TITLE Analysis of the chromosome sequence of the Legume symbiont Sinorhizobium meliloti strain 1021
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (17), 9877-9882 (2001)
MEDLINE 21396507
PUBMED 11481430

REFERENCE 2 (bases 1 to 333800)
AUTHORS Gouzy, J.
TITLE Direct Submission
JOURNAL Submitted (26-JUL-2001) Gouzy J., Submitted on behalf of the MELILO EU Consortium

COMMENT

MELILO EU Consortium:
Laboratoire de Biologie Moléculaire des Relations Plantes-Microorganismes, UMR215-CNRS-INRA, BP27, F-31326 Castanet, France, Laboratoire de Génétique et Développement UMR061-CNRS, Faculté de Médecine, 2 avenue du Pr. Leon Bernard, F-35043 Rennes, France, GATC GmbH, Fritz-Arnold-Str. 23, D-78467 Konstanz, Germany, Universitaet Bielefeld, Biologie IV (Genetik) Universitaetstr 25, D-33615 Bielefeld, Germany, Unité de Biochimie physiologique, Université Catholique de Louvain, Place Croix du Sud 2, Bte 20, B-1348 Louvain-la-Neuve, Belgium, Unité de Microbiologie, Faculté des Sciences Agronomiques de Gembloux, Avenue Marechal Juin 6, B-5030 Gembloux, Belgium. E-mail: jerome.gouzy@unilouvain.be; jfr http://sequence.louvain.be/meliloti.html.

FEATURES

source

1.333800
/organism="Sinorhizobium meliloti"
/strain="1021"
/db_xref="taxon:382"
complement(233..1114)
/gene="SMC04043"
complement(233..1114)
/gene="SMC04043"
complement(233..1114)
/gene="SMC04043"
/function="miscellaneous; hypothetical; global homology"
/note="Product confidence: hypothetical
Gene name confidence: hypothetical
predicted by Codon usage
predicted by Homology
predicted by Framed"
/codon_start=1

CDS

gene

/evidence=not_experimental
 /transl_table=11
 /product="CONSERVED HYPOTHETICAL PROTEIN"
 /protein_id="CAC47405.1"
 /db_xref="GI:15075851"
 /db_xref="SPTREMBL:O92M30"
 /translation="MGEAVEMFEFFVEFPASQRIPIPVFNSEHSGRHYPQFFDOSRID
 PHSTRSEDDHFDLEFSATFGLGAPLLRAHPRAFDLVNREPELDMEDMGLAPPA
 NISMRYAGLGVPLRVLAENMEIYRPFVEQALERIEIYKPYHATLKRLARTHY
 EFGMSILIDCHSMGNVHLPGSGHRPDIIGDRGTSAAELSRAVELLEOLGYAAV
 RNKPYAGGFTEHGRPTRGHLAHLQIINGLYVDENLILKKPFALMEDLATFIA
 LARHVEDGAYLPLAE"
 1305.1.1667
 /gene="SMC04044"
 1305.1.1667
 /gene="SMC04044"
 /function="miscellaneous: not classified regulator"
 /note="Product confidence : putative
 Gene name confidence : hypothetical
 predicted by Codon_usage
 predicted by Homology
 predicted by Framed"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="PUTATIVE 2-COMPONENT RECEIVER DOMAIN PROTEIN"
 /protein_id="CAC47406.1"
 /db_xref="GI:15075852"
 /db_xref="SPTREMBL:O92M29"
 /translation="MTRKIIIAEDNDMRFLYKALEKAGKYIYSTNGASAYDLRE
 EPFSLITDIAVPEMDGIELARRATELDPOLKWFITGFVAVALNPDSKAPKQKAVLS
 KPFRLDLVNEVKMMLA"
 1952.1.2026
 /gene="tRNA-VAL_GAC"
 1952.1.2026
 /gene="tRNA-VAL_GAC"
 /product="tRNA-Val"
 /note="codon recognized: GUC; predicted by tRNAscan-SE"
 /evidence=not_experimental
 2459.1.4297
 /gene="tLYD2 OR SMC04045"
 2459.1.4297
 /gene="tLYD2 OR SMC04045"
 /EC_number="4.2.1.9"
 /function="small molecule metabolism; amino acid
 biosynthesis; isoleucine/valine"
 /note="Product confidence : probable
 Gene name confidence : probable
 predicted by Codon_usage
 predicted by Homology
 predicted by Framed"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="PROBABLE DIHYDROXY-ACID DEHYDRATASE PROTEIN"
 /protein_id="CAC47407.1"
 /db_xref="GI:15075853"
 /db_xref="SPTREMBL:O92M28"
 /translation="MPAYRSRTTGHGNNMAGARGLMRTGKMSDECKPIIAVNSPT
 QFVGHVHLKIDGLVAREIEAGVAKKEFTIAVDGIMAGDMGLYSIPSEIIAD
 SVEYMNNAHADAMVCISNCKITTPGMLAALRLNIPAFVSSGPMRAGKYVILGKTH
 ALDVNMAVAADKVSDEVDQIITERSACPTGSCSGMPTANSNCTETELGLSLPN
 GSTLATVHADKRLFEAGHLIDLARVYEQEDERYLPNIATKQAFENMALDIANG
 GSTNVLHLITAAVEGIEDFTMDIDRLSRKVPCLSVAPAKADVHEDVHRAADIAG
 ILGELDKGLINDCEPTVHAETLGDIDRMDITRTSDTVRKFFRAAGGIPPOVAS
 QEARMEIDTRENGLVRSVEHPEPSKDGGLAVLGNALDGCIVKTAGVNESILKTSQ
 PARVESQAAVVGILANETKADGVYIRKEGPKGPGMCEMLYPTSYLSKGLGRAC
 ALTTDGRFSGTISLIGHVSPRANGTIGLVREGMDIDIPNRTISLRVDEAELA
 ARTEQDAKMKPVEORKRRTTALKAFAFASADBGAVRDLGDR"
 4343.1.4400
 /note="Sm-5 OR SMC04615
 REPEAT SM-5
 predicted by Homology"

/evidence=not_experimental
 complement(4470.4874)
 /gene="SMC04046"
 complement(4470.4874)
 /gene="SMC04046"
 /function="miscellaneous: hypothetical/global homology"
 /note="Product confidence : hypothetical
 Gene name confidence : hypothetical
 predicted by Codon_usage
 predicted by Framed"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="CONSERVED HYPOTHETICAL PROTEIN"
 /protein_id="CAC47408.1"
 /db_xref="GI:15075854"
 /db_xref="SPTREMBL:O92M27"
 /translation="MKTHGSCGRIREVVIDLEAGTSRCNCYSKLRWGSAY
 KPDEFRLMCEAGIGDYQFGTMSGHRFCYACGVTPGHGYVEIGAFSINVACID
 GVPDAEFALPIQYMDGLHNNMNAPETRHM"
 4979.5024
 /note="REP (repetitive extragenic palindromic) element;
 predicted by Homology"
 /evidence=not_experimental
 5027.5084
 /note="Sm-4 OR SMC04584
 REPEAT SM-4
 predicted by Homology"
 /evidence=not_experimental
 5094.5542
 /note="Sm-2 OR SMC04873
 REPEAT SM-2
 predicted by Homology"
 /evidence=not_experimental
 complement(5586.6029)
 /gene="azuz2 OR SMC04047"
 complement(5586.6029)
 /gene="azuz2 OR SMC04047"
 /function="small molecule metabolism; energy transfer;
 electron transport"
 /note="Product confidence : probable
 Gene name confidence : putative
 predicted by Codon_usage
 predicted by Homology
 predicted by Framed"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="PROBABLE PSEUDOAZURIN (BLUE COPPER PROTEIN)"
 /protein_id="CAC47409.1"
 /db_xref="GI:15075855"
 /db_xref="SPTREMBL:O92M26"
 /translation="MKTKMLIAMLCATSAGOANAEERYEMLNKAADGVMAPFPAY
 IRAQPGTVYFAKDKDGNLSALMKGAPBEARWTKGINETVTLSPKQVYMYQCAP
 HVGGMIGALIVEPEANLEAVKIGIKTPGSKAAAEKIFATESG"
 complement(6076.6417)
 /gene="SMC04048"
 complement(6076.6417)
 /gene="SMC04048"
 /function="small molecule metabolism; energy transfer;
 electron transport"
 /note="Product confidence : putative
 Gene name confidence : hypothetical
 predicted by Codon_usage
 predicted by Framed"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="PUTATIVE CYTOCHROME C PROTEIN"
 /protein_id="CAC47410.1"
 /db_xref="GI:15075856"
 /db_xref="SPTREMBL:O92M25"
 /translation="MTHASSRIAAVILCVTIGLSPPMASAEEDKIALGREIFLSE

Query Match	40.7%;	Score 480.6;	DB 1;	Length 333800;
Best Local Similarity	64.5%;	Pred. No. 6.7e-46;		
Matches 717; Conservative	0;	Mismatches 394;	Indels 0;	Gaps 0;

1 ATGCGAGCCGTCGATCTTTGGGGGCGACGGGCTCCATCGGGGAATCCACACTTCGACCTC 60
 189367 AAGGCGCGTGTGACGATATTGGGCTCCAGGATCGATCGATCGGAGCACACAGCGCTCGATG 189426
 .61 GTCATCGGAAGGGCGGGGCCGAGAGCGCTTCGACACGTCGCTTGATCGACGGCGGGCCCAAC 120
 189427 ATCGAGCGCCCTCGGCGGAGCCGATCGTTTCGATGATCGCCGCTCGAACCGGCAAGGCAAT 189486
 121 ATCCGGGACTGCGCCGAATGCGGCGTGGGCTGAAGCGGAGCTTGGCCGTACCCGGCAT 180
 189487 ATACCGGTGTGGCCGAGGAGGGCGCGCGCATGGCGCCGACGTCGCGATACCGCGGAC 189546
 181 GAGACATCGCTGCGCCGCGCTGCGCGAGGCGCTGAGCGGAGCGGAGCAACCGAGTGTGGGGC 240
 189547 GAGGATGTGCTATGGCGGAGCTCAAGATAGGCTCTAGCGCGAGCGGCATCAGAGTGTGGCGCC 189606
 241 GGGGCGCAGGCCATCGCCGAGGGCGCGGACCGCGCGCGCGCATGTGCGGCATC 300
 189607 GGACGAAGCGGCGCTCACCGAAGGCGCGGACCGGGATCCGCGTGGGTATAGCGCGGCATC 189666
 301 GTGGGGCCCGGGGCGCTCGAGCCCGGAATCGGGGCGGTGAAGCAAGCGGCCGACAGCGTGGCG 360
 189667 GTCCGCAATGCCGCTCTGCGCCACCTTGGCGGCTGCCCCGCGGCGCAGACATCGG 189726
 361 CTCGCAACAAGAAAGACCTCTGTAGCGGCGCACTCTCTGTATGGGAGCGGCCACAGAG 420
 189727 CTTGCAACAAGAAAGATCTCTGTATCCGCGGCGAGCCTTTTCAATGATGCGCGTGGCGAG 189786
 421 AAGCGCGCAGATCTCTGCGGCGGTGGAGAGGAGCAGCTCCCGGCTTTTCAAGCGCTGGG 480
 189787 GCGCGCGCGGTTGGTGGCCCGTGCAGAGGAAACATCGATCTTCCAGGTGGTGGAA 189846
 481 GGGCAGGACACGCGCTCTGTGAGCGCGTATATCAAGCGCTGCGCGGCGCGCTTCGCG 540
 189847 AAGCATCAGGGCCACGCGCGTGGAAAGCATTTGTCTCACGCGCTTCGCGGCGCTTCGCG 189906
 541 GACTGAGCCTCGAGCGCATCCGCGCTGCACCGTGGCGGAGCGCGACGCGCATCCCAAC 600
 189907 ACCAAGACGCTCGCAAGAAATGCGGACAGTACCGCGGATGTGGCGCGCCCATCCGCAAC 189966
 601 TGGTCCATGGGCGCAGGGATCTCATGACAGGCGCTCGATTTCAACAAGGCGGTGAG 660
 189967 TGGTGCATGGGCTCCTCAAGATATGATGACAGGCGCTCGATGTTTCAACAAGGCGGTGAG 190026
 661 CTGATCGAGAGCGCGAATCTTTCGCTTCGAGCGCGGACCGGATTCGAGCGGTGTCTCAT 720
 190027 ATGATCGAGGGCGGCCACCTCTTCGCTTCGAGCGCGGAGCAATTCGATGTGTCTCAT 190086
 721 CCGCATTCATCTGTCAATGCGATGTGGCTTCTGCGAGGGGCGCTGATGSCCATCTC 780
 190087 CCGCAATCGGTGTCTACTCTCATGTGTGCGCTAATACGACGCGTCCGACTCGCACAGCTC 190146
 781 GGGCGCGCGCATGGGCGCAAGCGCATGGATGGCGTGAATCGGCGGGGTCCGGCGGAG 840
 190147 GGGCTGCCCCGACATGCGCACCGCATGGCTATGCCCCGTGCTTCAACCGGAACCTCGGAC 190206
 841 GTGCGCTCGCGCGGATCGACTTCGACACATTTGCGAGCTTCACCTTCCAGAACCTGAC 900
 190207 CTGCGCGGTGAGGGCTCTGACTTCGACAGGCTGCGCCGCTTATTTTCAGAGGGCGGGAC 190266
 901 GAGGACGCTTTCGCGCCCTGAGGCTTGGCGGAGACGTCAATGGCGCGCGCGGCTGTG 960
 190267 GAGGTGCGCTTCCGCGGATCAAACTCGCGCGCGGCAATGAGAGGCGCGCGTGCAG 190326
 961 GGGCGCGCTTCAACGGGCGCAAGGAGATGGCGTGCATCATTTTCAATCGCGGACGATC 1020
 190327 GGGCGGGTGTCTGAACGGCGGCCACAGAGACCGGCTCTGATCTCTTTCATCAAGGGCGCATC 190386

[illegible]

Db 190447 GCCGCCACGAGCATGGACGACGCTCTTGGCG 190477

RESULT 4
AE008174

DEFINITION

Agrobacterium tumefaciens str. C58 circular chromosome, section 232 of 254 of the complete sequence.

VERSION
KEYWORDS
SOURCE

Xerobacterium tumefaciens strain C58 (Gordon)
AE008174.1 GI:15157805

ORGANISM

Agrobacterium tumefaciens str. C58 (celeon)
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group; Rhizobium.
Rhizobiaceae: Rhizobium.

REFERENCE
AUTHORS
TITLE

1 (bases 1 to 520)
Hinkle, G., Slater, S. C. and Goodner, B.
Complete Genome Sequence of *Aerobacterium tumefaciens* C58

JOURNAL

Unpublished

AUTHORS
TITLE

Hinkle, G., Slater, S.C. and Goodner, B.
Direct Submission

FEATURES

Cambridge, MA 02139, USA

```

/organism="Agrobacterium tumefaciens str. C58 (Cereon)"
/strain="C58"
/ab_xref="taxon:191661"

```

gene
CDS

```

229 . 981
/ gene="AGR_C_4722"
229 . 981

```

```

/ncbi="ncbi_0004722"
/note="probable transcription regulator PA3757 (imported)
- Pseudomonas aeruginosa (strain PA01)"

```

```

/transl_table=11
/product="AGR_C_4722p"

```

```

/db_xref="GI:15157806"
/translation="MNGALINLEDLQSGGPLYLKLROTIEDAINGRKLKHGDALPEP

```

RRKGMTASTQWIDRGLFHPGTDEMTLGLSQSAMVARLTRRLANDQPIALEYTSLPE
DILPDQVENSPLYLEKSGIRPVRAIQIRISARNLTDEETLLGVPQGSALTSVQRR

gene

```
1006. : 2028
/ gene="AGR_C_4724"
# 2000
```

/gene="AGR_C_4724"
/note="glucosamine--fructose-6-phosphate
metabolic pathway related to D-glucose metabolism"

```
(strain kl)"
/codon_start=1
/trancl_table=11
```

```

/product="AGK_C_4/24"
/protein_id="AAK88329.1"
/db_xref="GI:15157807"

```

IVALLAOSARGGALLTESLVNTLPPSPLEGAADHAINIOAGPEKAVAATKSFVNISVAGCI
IARGSSDHAHFLKYAIELOGTGLPVASGLPSLASIYOTKRLRLERAARAFVAVSQSGSPED
/TANSLAHLHMI LUNGLVLELLE DAVANNDMMNTRILREKALANLZEVAV
TANSLAHLHMI LUNGLVLELLE DAVANNDMMNTRILREKALANLZEVAV

AAALKKECEELHAEAYSSAEEMHGVSVLAPQEPVIAFAARDRAETSVTDIAASLANNN
GANEVYSTKGSAPKPLPFVETNHP L TDALLIAPFYGFIEOLSRARGFNPDAPVALL

gene

2025, .3191

gene
 CDS
 /gene="AGR_C.4726"
 20251..3191
 /gene="AGR_C.4726"
 /note="N-acetylglucosamine-6-phosphate deacetylase Xf1465 (imported) - Xylella fastidiosa (strain 9a3c)"
 /codon_start=-1
 /transl_table=11
 /product="AGR_C.4726p"
 /protein_id="AK88330.1"
 /db_xref="GI:15157808"
 /translation="MSGIKAFVCAKRFIDCAAMEGRVYVGDGHVTAISDVPAEAM
 IDVKGILIAAGETIDLVNNGGVWFMNNQDVEGIARICSHAFRTALVTLITLPR
 DIVSKAAQAGIDSKOVPGFGLHFGPGLSVARKGTHDPALVRKMETADLVLLGC
 KALFAVMTTIAPEYNEEQVVALRAGIVSGHTDGLDVAIVTAEGASVTHLF
 NAMPPLGHPREPVGALNSNGDCGLIADGFPVDPALIGIALRAKPGPRTLVDA
 MSTIGTDDGFELNGRRVYNGRGLTLEDGLAGADIDMLSCVRPMYEXKEMPLEEL
 RMASAPPAVAGVSDKSKLLPFPDAPFVLTLPOLQHMSTIGSEKTYDAVHSV"
 3195..4121
 /gene="AGR_C.4728"
 3195..4121
 /gene="AGR_C.4728"
 /note="ROK family protein VC1532 (imported) - Vibrio
 cholerae (group O1 strain N16961)"
 /codon_start=-1
 /transl_table=11
 /product="AGR_C.4728p"
 /protein_id="AK88331.1"
 /db_xref="GI:15157809"
 /translation="MTVCPDIGGTTIKGALAHPPDDIRPVPIPTPKTSFEDPAAGIK
 SVTDAAGGTPGCYSLSIAGYIDPDTGKATYANPSIHGRYLKDELEKALNLPYVSND
 ADCFVAIESIEIGGGGRVYEVYILGCGGGLYIDCKLINSHGFEAGEHCPVAT
 SAGNPVSLREFECGGGLTGCVDAGISARKEMLHAHLHROEMTSEDIITAAWQAGAK
 AARTIVLDIILASPLAMVINTGATIVPGGLSRELLALDEAVNRGRIIRFRNR
 PIVPAICRIEPLIGSAVIGLKRYQEIARPS"
 complement(4128..5399)
 /gene="AGR_C.4730"
 complement(4128..5399)
 /gene="AGR_C.4730"
 /note="probable hydoxylase - Mycobacterium tuberculosis
 (strain H37RV)"
 /codon_start=-1
 /transl_table=11
 /product="AGR_C.4730p"
 /protein_id="AK88332.1"
 /db_xref="GI:15157810"
 /translation="MGARVMEADAKAGHEKDRVAVONITHCVAAGGGPACMGLLIA
 RSGIDVYVEKHGDFLDFERDITHSSTLEMLQLEFIDEFLSLPHRAFLNAVIGG
 ERYTIDFESRLPYRHRPIAPMOMDLNFTVGAGOVENRILMANAPVGLIERDGV
 GGLTVEPCEVITADIYVAGDRNSIYEAGLEVQRCGTFTEVLMLELSQPDPP
 TETMGAGSRQGVMIINRDYVQCGTVYKRGFTADIKLELEAFRTVAETICFPBR
 LDELKSMDDVHLTVRIIDRLKRWKPGLICIGDAAHAMSPGVGNLAIDAVAAPAN
 ILVPLVSGRAIADDLALAEKRSPPTKATQFLORMMRIGRRKQGEERAEKPKGPPA
 FVRAIIRFPLSLHTGLVGLGFRREKIEPR"
 complement(5419..6072)
 /gene="AGR_C.4732"
 complement(5419..6072)
 /gene="AGR_C.4732"
 /note="hypothetical protein PA4011 (imported) -
 Pseudomonas aeruginosa (strain PA01)"
 /codon_start=-1
 /transl_table=11
 /product="AGR_C.4732p"
 /protein_id="AK88333.1"
 /db_xref="GI:15157811"
 /translation="MLVLFIPDPRLSETAOLPGRIYVFNRAITDFGTFRMMLYST
 GALAIAYVAARVLOAQOTYGGRLTAMRIILAFPLITGTSTIIVHTLKLIGARPBL
 FLEMGAYSLEPFTGDNLYESFSGHSTAGAGFCVRAMLPRIIRGPFLLALVATGR
 VIVGAHPSDVAAGLLDGLWTAMAFATIRAFBMLFRPDMQGNPQPKNAPSAAETR"
 6251..7444
 /gene="AGR_C.4736"
 6251..7444
 /gene="AGR_C.4736"
 /note="1-D-BOXY-D-XYLULOSE 5-PHOSPHATE REDUCTOISOMERASE

(DXP REDUCTOISOMERASE)"
 /codon_start=-1
 /transl_table=11
 /product="AGR_C.4736p"
 /protein_id="AK88334.1"
 /db_xref="GI:15157812"
 /translation="MTNASEMPKRLITLGSIGTNTLDVYRQLOGRDEFEIMALTG
 AGNIALIAEQAREGAOLAVTADDYKLSALACTGILVAAAGAGLEEAASDAGM
 VMAIAGTGPALPTLTAARRGADIALANKCLISADVFLRTKOGGGRLIPVDSHS
 ATRQCLGEGKOAVERIVLTASGGSPRTSRDMSVNTADIAAHNMWSGLKVSIGS
 ASMFNKGLEMTIEKYLEPDLRPDYDVIYHPOSTIHSNVTGDSYTAOLGSPMRRI
 SVALTTPERNLSVERIDFAKLRLDFEADDEARFPALRLARALERGGQALANA
 EETAFHAFVAGGIGFLDMAEIVTEVMDRMDHGTAEFTMDVFSADDEARRHALELAT
 KERAA"
 complement(7480..8751)
 /gene="AGR_C.4738"
 complement(7480..8751)
 /gene="AGR_C.4738"
 /note="5-AMINOLEVULINATE ACID SYNTHASE
 (DELTA-AMINOLEVULINATE SYNTHASE) (DELTA-ALA SYNTHETASE)"
 /codon_start=-1
 /transl_table=11
 /product="AGR_C.4738p"
 /protein_id="AK88335.1"
 /db_xref="GI:15157813"
 /translation="MTVPYSERKTVSITLHGMDFEAFETTELQSLHSEGRVYFADI
 EROGNEPRATRYNANGDRKDYVWCNSNDLGGOMPKVTEAKKAIDHCGAGAGTR
 NISGTNHYHVLEOELADLHGKESALIFTSYVSNMKTGLGKLIIFSDALNH
 ASMIEGIRVGRCEVIMKNDLHDLAKIKAADPNPKLIAFESVYSMOGDIAPIREI
 CDLADRGANTYIDDEVYHAGMYPRGGGIAERGLMDRLITIGTIGKARGVGGVIT
 GSTAVCDYFISFASGFTFTTALPPSLAAGIASIOHLKASPFERARHODVRKLRGL
 DAGIIPMDNPSHIVPMVMDAKCKWISDILDNHGVYQIPNTYPTVPRKTERLRT
 PFLIHDADIEQLVGLAHLQMSICALARAVA"
 9017..9319
 /gene="AGR_C.4740"
 9017..9319
 /gene="AGR_C.4740"
 /note="hypothetical protein"
 /codon_start=-1
 /transl_table=11
 /product="AGR_C.4740p"
 /protein_id="AK88336.1"
 /db_xref="GI:15157814"
 /translation="MGRLPERRAHCVNSESFMKKAIIFALVGLASCTOTEGASIG
 AVSGAIVGGAVTNRGAAGAAGIAGGALIGNASEPGCYRDOYRGORYTARCR"
 BASE COUNT 1974 a 2842 c 2765 g 1855 t
 ORIGIN
 Query Match 36.7%; Score 434.2; DB 1; Length 9436;
 Best Local Similarity 61.9%; Pred. No. 2.0e-40;
 Matches 688; Conservative 0; Mismatches 423; Indels 0; Gaps 0;
 QY 4 CGCAGCCTGTCGATCTTTGGGGCCACCGGCTCCATCGCGCAATCCACTTCGACCTGTC 63
 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
 Db 6275 CGCAAGCTGACCATATTTGGTTTCAGCGGTTCCATCGGCACCAATFACCTGAGATGTCG 6334
 QY 64 ATGCGGAAGGGCGGGCCGAGCGCTTCGCAACGCTCGCTTGACCGCGCGGCGCAACATC 123
 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
 Db 6335 CGTCAACCTGGGTGGCCGCGAGCGGTTGCAGATCATGCGCTGACCGCGCGGGAACATC 6394
 QY 124 CGCGGACTGGCCGAATATGCGCGCTGCAAGCGGAGAGCTTGGCGTCCACCGGCATGAG 183
 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
 Db 6395 GCGCTTCTGGCGCAAGCGCGGGGAATTTGGCGCGAGCTCGGGGTCTACCTCGCGAAGAT 6454
 QY 184 GACTGCTGCGCGGCTGCGCGGCGCTGCGCGGACCGGACCGAGAGTCCGCGGCGG 243
 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
 Db 6455 GATTAATATGAAAGCGCTGAAATCGGCGCTGCGAGGACACCGCATCAAAAGTCCGCGAGGC 6514
 QY 244 GCGCAGGCGCATCGCGCGAGCGCGGACCGCGGCGCACTGAGACCATGTCGCGCATGCTG 303
 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
 Db 6515 ATGCGCGGCGCTTGAAGAGCGAGCGCTGCATGATGATCGCGGTGGATGATGCGCGCATTTGCC 6574
 QY 304 GCGCGCGCGGCGCTGCTGCGCGGAGATCGCGGCGCTGAAGCAGCGCGCAGCTGCGCTG 363

Db 947 AGCGTGGAGCGCTGATTTCCGCAAACTGCGACGCTCATTTCCAGACCGGAGAA 1006
QY 904 GAGCGTTTCGGGCGCTGAGGCTTGGCGAGATCATGGCGGGCGCTGCGGGC 963
Db 1007 GCGCGTTCCCGCGCTGCGCTTGGCGAGCGCTGAGCGGCTGCGGGC 1066
QY 964 GCGCGCTTACGCGCGCAAGAGATCGCGCTCATTTCCATTCGCGCGGAGCATCGGG 1023
Db 1067 GCGCGCTTACGCGCGCAAGAGATCGCGCTCATTTCCATTCGCGCGGAGCATCGGG 1126
QY 1024 TTTCTGACATGCGCGCTGCTGCGAGAGACGCTCGCGGCGCTTTCGACGACCCCTG 1083
Db 1127 TTTCTGACATGCGCGCAAAATGTTGAACGCGTATGATGACGCGCGGAGC 1186
QY 1084 TTTCGAAAGTGGCGCGCGCTTTCGAGGAG 1114
Db 1187 GCGGAGAGCATGAGCGACGCTTTTTCGCGAG 1217

RESULT 6
AE005865/c 12977 bp DNA linear BCT 12-JUN-2002
LOCUS Caulobacter crescentus CB15 section 191 of 359 of the complete
DEFINITION genome.
ACCESSION AE005865 AE005673
VERSION AE005865.1 GI:13423361
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
1 (bases 1 to 12977)
Nietman, W.C., Feldblyum, T.V., Laub, M.T., Paulsen, I.T., Nelson, K.E.,
Eisen, J., Heidelberg, J.F., Alley, M.R., Ohta, N., Maddock, J.R.,
Potocka, I., Nelson, W.C., Newton, A., Stephens, C., Phadke, N.D.,
Ely, B., Deboy, R.T., Dodson, R.J., Durkin, A.S., Gwinn, M.L.,
Haf, D.H., Kolonay, J.F., Smit, J., Craven, M.B., Khouri, H.,
Shetty, J., Berry, K., Uterback, T., Tran, K., Wolf, A., Vamathevan, J.,
Ermlaeva, M., White, O., Salzberg, S.L., Venter, J.C., Shapiro, L. and
Fraser, C.M.

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
Complete genome sequence of *Caulobacter crescentus*
Proc. Natl. Acad. Sci. U.S.A. 98 (7), 4136-4141 (2001)
2 (bases 1 to 12977)
11259647

Nietman, W.C., Feldblyum, T.V., Paulsen, I.T., Nelson, K.E., Eisen, J.,
Heidelberg, J.F., Alley, M.R., Ohta, N., Maddock, J.R., Potocka, I.,
Nelson, W.C., Newton, A., Stephens, C., Phadke, N.D., Ely, B.,
Laub, M.T., Deboy, R.T., Dodson, R.J., Durkin, A.S., Gwinn, M.L.,
Haf, D.H., Kolonay, J.F., Smit, J., Craven, M., Khouri, H., Shetty, J.,
Berry, K., Uterback, T., Tran, K., Wolf, A., Vamathevan, J.,
Ermlaeva, M., White, O., Salzberg, S.L., Shapiro, L., Venter, J.C. and
Fraser, C.M.

Institute for Genomic Research, 9712
D 20850, USA

04:26 2003

protein family: HM

/translation="MTINRDKPLKYLVAEASGALGAGLAKALRARLGGVTEVGI
GAKMAEQVSPFIDIALSLIIGLESLKAYRPARARLKDVALAAREKPDVALIDS
WGENFALARRLPDLPFLVYVPOVWYARQALAKAVDLISLOPKDYFD
AGLIONFVFNALAKRDPHADPGRARRAIGAAASOILLVPGSPSEIEVMFAE
DAVRLKVRPDLHIVPAATVVAEYKARVAPFRHAYIDECKDAFLAVAPAE
ACGVITTELALAGRPVGVYGTAVTAIVRLKMPRTITLNTLAAGAAVAPLLOH
ACEGBGLAREVALRLDDPDLRQRTAEYVALDRMGKMPDESEAAALDLFLSARG
ALGRAG"
/complement(1267..2109)
/gene="CC1910"
/note="Identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAK23885.1"
/db_xref="GI:13423363"
/translation="MRKGLIAGGALPYELASHCEAAGRAVNRRLRSFADPSIDRY
PCADVIGEFGRKIFKALRAEGCDVYCFAGNVRSPDSEALPARGLKPLSLVARK
GPDALLRLVDEFEKRGFEIRGAEHVEGEMTLPRGIRGVRAPRHMADIDKLVAR
EIGRLDIGGAYVCSGLVLAVERQESTDMALRVADLPRAIRGRARRLGVAKAPK
IDETRVLDPTIGVATLIRARAGLAGIVGEAGRLLVDEEAVIAADDLGLVGLVDP
QERP"
/complement(2119..2910)
/gene="CC1911"
/complement(2119..2910)
/gene="CC1911"
/note="Identified by match to protein family HM"
/codon_start=1
/transl_table=11
/product="acyl-(acyl-carrier-protein)--UDP-N-
acetylglucosamine O-acyltransferase"
/protein_id="AAK23886.1"
/db_xref="GI:13423364"
/translation="MSINPPTAIIAPAKLAPDEVGSPISVDPVTLAGVRLSHV
VCAATIGGCVVSHFANLGPPOHGHGERTELIGRNIRIHEVMTHTASGK
VTTIGSDGLKVMGSHVAHDCVTGDFVLAKKATIGGVAIGVYEGGLAAHOFERI
GRSFGLAATVATKDYIPIGYSWGNNAHEGLNVLKRGPPREINLRLALME
ADECTPQERLDVAETHAGNAEMELVIDIRIDANRPLCLPERV"
/complement(2907..3386)
/gene="CC1912"
/complement(2907..3386)
/gene="CC1912"
/note="Identified by match to protein family HM"
/codon_start=1
/transl_table=11
/product="(3R)-hydroxymyristoyl-(acyl-carrier-protein)
dehydratase"
/protein_id="AAK23887.1"
/db_xref="GI:13423365"
/translation="MGDNAEQAVNDIDIAELIARIPHRYPPLVNDARDYNPHOSIV
GIKVTINPEFPOGHFPGNPPGVILITLALQTCGLVMSKSLVDTEKTIFFKSYD
NAKFNPPRRPGDVRMEVEVLRARSSIFKGVAKVGRVAAEAPAAVETGPRA"
/complement(3389..4408)
/gene="CC1913"
/complement(3389..4408)
/gene="CC1913"
/note="Identified by match to protein family HM"
/codon_start=1
/transl_table=11
/product="UDP-3-O-3-hydroxymyristoyl glucosamine
N-acyltransferase"
/protein_id="AAK23888.1"
/db_xref="GI:13423366"
/translation="MPDPRFDSLGPALLISELAQAATLADALGERVITTAAPLDA
SDAQATTFESDAKRDAASTRAGCFVPRHQGLPPICALLVYGRPAAPAAVAAANR
LHAPRRHAGASLHPDALEDGVALAPVITIGOGASIGRTIRIGPVVIGGVYGR
YCRIGANAVIGFMDGNVAISAGVIGAGFGAALGPRGVLDPLQIGVYVODNVT
GANSVGRGAFGPTTIGENTKIDNLTIVAHVHNYRIGENCYLAAYTGVSSTVVDGVA
GKAGVADHLNIGSGASIGAAASVFEVDVDEGTWICFPARPLKRLRLRETNALSRAAG
RGRG"
/complement(4462..5100)

gene

gene
CDS

...5100)
 /transl_table=1
 /product="conserved hypothetical protein"
 /protein_id="AAK23889.1"
 /db_xref="GI:13423367"
 /translation="MSKFLPAASSVVALAASSASNOTAPAPAPAVYTHGPAALPG
 VCIFSSPRAVSSLVGKADARLKTIIQVNAELTGERTALDNEAKALDAKKTIIAD
 ALFOQATLOAKANAMORCOLROKEVEAEFEOKALSHVOELNTPIOOYVOAKCSVL
 LDREAVLAMPANDIDYVAALDARKITLTFDRERLDQVPCAAALQPIPK"
 complement(5103..7412)
 /gene="CC1915"
 /complement(5103..7412)
 /note="identified by match to PFM protein family HMM
 PF0103"
 /codon_start=1
 /transl_table=1
 /product="outer membrane protein"
 /protein_id="AAK23890.1"
 /db_xref="GI:13423368"
 /translation="MALLGSTALVAPQQAQAOTGVVORITIVGNERIEGTVLS
 YLPDGPVDSORDLAKTLARFDLADVKLEMLGDLVYKVENPITNOVFEEN
 SSIKEDKLAEVOIRPRGIFTRAKVQADVORIELRRSGRISITVPKVELPQKEV
 DLVEINEGAKSVLGINFLGNAEYSDNDLDVIVTKESWYILTSNDVDDRLRY
 DRQLRHHYNNRGYDFEYVSSVAELAPDNGFAVYTLLEGPRKYGKTLITETELK
 LDENLAQILPRTGOLYEDERIEQATDALTFAAGAAFNVDYRPRYVNRRTKYD
 LVQVREGPRVYDRIQDYNTERLDVLRLELVEAGDYNVLYDRSKNNRRIGF
 KEVEIEDAGSAPDRSLRKVEEPOPTGELSPSAGYSDTKLVLDVQCTERRPRGQ
 QNRRARASVSLRQOIDFGESSEPRFLGRNLYAVNLYTPFDLSEFAYDTKVGQV
 REGPPLTNDSSMLRTRVRODEVSVADSLCASVSQIICLQCAVITSLIGGLRID
 KRNDIPNTRGMEADLNQDLAVGQDVKYDKTADAGWYGFRTKDVFSAGSGFYIE
 GMGCDNVRINDRFYRGSTSRGFEIAGIGRDISSEFSNGAKLYAISTEFLVPTFL
 PEQYGIKAALESVGTAGLLDQVDRQSPVFPNPKNDNLGNASGISIDKMSPKMP
 IRPDISRLSKKEYDRETFRSTSTRFQ"
 complement(7484..8680)
 /gene="CC1916"
 /complement(7484..8680)
 /note="identified by match to TIGR protein family HMM
 TIGR0054"
 /codon_start=1
 /transl_table=1
 /product="membrane-associated zinc metalloprotease,
 putative"
 /protein_id="AAK23891.1"
 /db_xref="GI:13423369"
 /translation="MTGFLIMLVSLFVLSVVVTHELHYAARACVAIERFSIGF
 GAPIISWRDRKGVENCVASIPLEGVRFADENAAVPPDNDLDAKNEIRREGDA
 VNRYFHRKPVMOAFIAVAGPANFIALIVFVIVISFQAGCTSTTVPKVELPQKEV
 AAGFKRGVGLKADNRQISDFDIOGVYLRANPIDEAVERGRVYHLATRLVER
 ONEISGRVYVGLGLRSAPGREGSSLSAIDEAVENDVKTIAFYGRLLMOL
 PADQISGIIIGHTAGAVTNGVYQAPNKAALIGILYSPFMIASLSISIGMNLIP
 IYVLDGHLVMTAYEAVAKRPLAEFQAGFRAGLALLIGFMLEAAMNDLNRVDERF
 IGGLEF"
 complement(8703..9902)
 /gene="CC1917"
 /complement(8703..9902)
 /note="identified by match to TIGR protein family HMM
 TIGR00243"
 /codon_start=1
 /transl_table=1
 /product="1-deoxy-D-xylulose 5-phosphate reductoisomerase"
 /protein_id="AAK23892.1"
 /db_xref="GI:13423370"
 /translation="MGALTSPRKVVVLGSGTIGLSTLSFEESGAPVOIALTAGRN
 VERLIOARMKPSLAVIEDESRDLRAGLAGTVEAAGADAVDAAMGADWYS
 AIVGAGLAPVYAAARTGAVIALANKESHVACGAPALLATAKAGGSGVTPDSESAIF
 OVLOSFCARHVSRLITLTSAGGPRRTDKAAMARATPEQALAHNMMSGAKISVDSATM

Query Match 29.8%; Score 352.6; DB 1; Length 12977;
 Best Local Similarity 59.0%; Pred. No. 3.4e-31;
 Matches 626; Conservative 0; Mismatches 429; Indels 6; Gaps 1;
 MKNGLEMIENASYLFATPEDEVYVYIHPOSTIHSLVYVDSSTIAQLGPPMRAPICAA
 FAMPDRLPMPARLDIAATGQLTFESDEPRFATIGRALRLGGAPRANMAAEV
 AVAFEDRRIRIGFLDIAGAVAGTALRNMNSLDLSSVASDAETMLIDGSAIRIAAEV
 QY 4 CGACGCTGTGATCTTTGGGGCCACGGCTCATCGCGCAATCCACTTCGACTGCTC 63
 Db 9881 CGCAAGGTGGTGGTGTAGGCTCAGCCGAGATCGACTGCGCTTTCGACTTGGCTCTTC 9822
 QY 64 ATGGCGAAGGGGGGGGGGGAGCGTTCGCCACGCTCTCGAGCGGGGGGCAATC 123
 Db 9821 GAAGAGTCGGGTGGC-----CGTTTCAGATCTTGGCGCTCACCGCGGCTCGCAATGTC 9768
 QY 124 CGGCGACTGCGCGAAATAGCGCTGCGCTGAAAGCGGAGAGCTTGGCGTACCGCGCATAGAG 183
 Db 9767 GAGCGCTGATCGAGCAAGCGCGGGCGCTGGAAGCGCTCCCTAGCGGTCATCGAAGACGGA 9708
 QY 184 GACTGCTGCGCGCTGCGCGAGCGCTGCGCGGACGGGACCGAGAGTCGGCGCGG 243
 Db 9707 AGCGGCTTGAATCTTCCGCGCGGCTGCGCGCTGCGCGGCTGCGAGAGCGCGCGCGG 9648
 QY 244 GCGGAGCGCATCGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 303
 Db 9647 GCGGAGCGGCTTCCGAGATCG 9588
 QY 304 GCGCGCGCGCGCTCGTGGCTC 363
 Db 9587 GCGCGCGCGCGCTCGCTG 9528
 QY 364 GCGCAAGAGCAAGCGCTGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 423
 Db 9527 GCGCAAGAGCAAGCGCTTGTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 9468
 QY 424 GCG 483
 Db 9467 GCG 9408
 QY 484 GAGGACAGCGCGCTGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 543
 Db 9407 GAGTGGCG 9348
 QY 544 TGAAGCTCGAGCG 603
 Db 9347 TGGGACAAAGCGCGCTATAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 9288
 QY 604 TCCATGGGCG 663
 Db 9287 TCGATGGGCG 9228
 QY 664 ATCGAGAGCG 723
 Db 9227 ATCGAGCG 9168
 QY 724 CAATCGATCGTCCATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 783
 Db 9167 CAATCGATCGTCCATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 9108
 QY 784 CCG 843
 Db 9107 CCG 9048
 QY 844 CCGGTCG 903
 Db 9047 CTGCG 8988
 QY 904 GAAGCGCTTTCG 963
 Db 8987 GAGCGGTTCCCG 8928
 QY 964 GCGGCTTCAACG 1023


```
/protein_id="AAW71371.1"
/db_xref="GI:21646023"
/translation="MEDKLSIKDKFQTIIEQQLSDPEVVADQNRFRKLKNEYSSSLKEI
VRAYDSMTKKOLDRAHSGKNNEDEMPALVEEAGELEQLERLPLEQQLKILLPLK
DEADSRNAIEIRAGTGDEAGLFAADLRMYOYAEROGWSQCTLVLEPSEVPGSLK
EVSLESGHNVYGIKAFESVHRVORPETEGRIHTSAASAVAPLEAEVDVEIRK
EDLIDTFRSGGKGNVKNKVERAVRITHVPSGIIVACQBEERSQLQNRERAMKLRK
LYDLQAEQKSRADLRSMVTTCGDRSAKIRTYNFPQSRVTDHRIGFTSHALPQIMQG
ELDPLEALRMHDQERLAETA"
/complement(5802..7163)
/gene="Ct0124"
/complement(5802..7163)
/gene="Ct0124"
/notes="identified by match to PFAM protein family HMM
PF00595"
/codon_start=1
/transl_table=11
/product="membrane-associated zinc metalloprotease,
putative"
/protein_id="AAW71372.1"
/db_xref="GI:21646024"
/translation="MELLNTIFFVIAFLVTAHEFHFIITARMFGMRVDRFFIGFD
FWGIKLWQKIGTEYIGAFPIGGVYKLAGMIDESMDTHVSOEVPWEFRAPVWQ
RLIYLAGGVAMNVLAAVIFIGTISFGESEPTITPAFIEPKSVFSSMGQSGDHLV
AIMGKALHWEERLDELPERLASGLQTIERNGEELTLTAPKDIISRINGNSIGIRPT
PVVIDQVLPGDPAARAGIMPGGLITAINSGPVADWSEVNIISANAGKLLVTHMHL
KNSTGPLLAALIRKAGQITTEVTNPSGKIGISLKQTIETIRIKLSLQALIASGLN
QWKTIVLVQVGGFKIFSKEDEPRKSVGGPKIARIANOSAEQGPISFMYFVAVLSIS
LATINILPALDGGQFVLNAIEGIMGREIPEVFKRIQOVGMVTLMLLFAYFMINDL
LNP"
/complement(7209..8357)
/gene="dxr"
/notes="Ct0125"
/complement(7209..8357)
/gene="dxr"
/notes="identified by match to PFAM protein family HMM
PF02670"
/codon_start=1
/transl_table=11
/product="1-deoxy-D-xylulose 5-phosphate reductoisomerase"
/protein_id="AAW71373.1"
/db_xref="GI:21646025"
/translation="MKSLSILGSTSIGSLTLDVYRRHPEPSIALAEHVDVEMLLK
QIDFSPISVVRDEASRLKGMGLDHPEILCGLGAEAAVADGMDVSAIVGA
AGLVPTVIAEAGKDLTALANKETLVVAGOLVSDLVKKHDKVLLPVDSEHSAIFQSLVG
HRTEDIERILTAGSGFPFRFAELKNVPEQALKHPQWSMGAKTIDSATLMNKG
EVTEAHLDPMAEKIGVYVHPOSIIHSMVEYIDGCVIADLGVPDMRAPYALAWPE
RCTGIGKLDLTKVATLTPEEDPMERFALRLAFADKAGOTYPAVLNAANEIATAAF
LDKIGFTDIAGTVDTQMHEAWTPTILEEYLQADKWARQILIG"
/8515..8591
/gene="trnA-Val-2"
/8515..8591
/gene="trnA-Val-2"
/product="trnA-Val"
/complement(8634..8912)
/gene="Ct0126"
/complement(8634..8912)
/notes="identified by match to PFAM protein family HMM
PF00708"
/codon_start=1
/transl_table=11
/product="acylphosphatase"
/protein_id="AAW71374.1"
```

```
Query Match 25.8%; Score 305.4; DB 1; Length 11150;
Best Local Similarity 57.2%; Pred. No. 6.8e-26;
Matches 602; Conservative 0; Mismatches 436; Indels 15; Gaps 2;

QY 9 CCTGTGCATCTTTGGGCCACCGGCTCCATCGCGGAATCCACCTTCGACCTCGTCATGCG 68
Db 8349 CTTATCCATTCTCGGCAGTACCGGCTCCATTGGACTCAGCACACTCGACCTCGTCAGCG 8290
```

```
QY 69 -----GAAGGGCGGGCGCGAGGGCTTCGGCACCGCTCGCTCTGACCGCGGGCGGCACACAT 122
Db 8389 TCATCTCTGAAGGTTTTCGATTGCGGCCCTTTCGCGAGGGCCACGACCTCGAAATGCTTCT 8230
QY 123 CCGCGCACTGGCCCAAAATGGCGGTGCGCTGAAGCGGGAGCTTCCCGTCAACCGGCGATGA 182
Db 8229 CAGCAAAATCGATGAATTCAGGCGGTCTACTGGTCTCGGTGCGGTACGAGGGGTCCGGGA 8170
QY 183 GGAATGCTGCCCGCGCTGCGCGAGGCGCTGGCGGGACGCGGACCGGAGGTTCGCGGGCGG 242
Db 8169 GCGCCTCAAGGGGATGCTCGCGCATCAAGCCGGAATTCCTCTCGCGCCTCGAAGGTGC 8110
QY 243 GGCACAGGCATCCCGAGGCGCGCGGACCGCGCGGCGGCGACTGGACCATGTGCGGCGCATCGT 302
Db 8109 AGC-----TGAAGTTGCGCGCTGATGGAGCAGACATGTAGTTTTCGGCGCATCGT 8059
QY 303 GGGCGCGCGGGCTCTGTCGCCGAATCGCGGCGCTGAAGCACGCGCGCAGCTGGCGCT 362
Db 8058 GGGTGGCGCAGGACTGGTGCCAAACCGCTCAGGGCTATTGAAGCCGGAAGGACATCGCGCT 7999
QY 363 CGCCAAACAAGAAAGCCTCGTGCAGCGGCGCAACTCTCTGATGCGGACGGCCCGAGGAA 422
Db 7998 GGCACAAAGAGAGCGCTCGTCTGCGCGGTACGCTGTTTCGACCTCGTCAAAAGCA 7939
QY 423 CCGCGCCACGATCCTGCGGTGGACAGCAGCAGCTCCGCGGTCTTTTCAGGCGCTGGCGGG 482
Db 7938 CGATGTCAAACTGCTGCGAGTGGACAGCAGCAGCTCGCGCATCTTTTCAGTCTGCTGCTCG 7879
QY 483 CGAGGACACGCGCTCGCTGCGAGCGGTTCATCATCAGCGCGCTCCGCGGGCGGTTCGCGGA 542
Db 7878 CCACCGCACGGAAGATATCAAGACGATCATCTGACCGCATCAGGTGGCCCATTCCTCGTAA 7819
QY 543 CTGAGGCTCTGAGCGCATTCGCGGCTGCACCGTGGCGGAGCGCAGCGCCATCCCAACTG 602
Db 7818 AACGCCCGCGAAGAGTTGAANAACGTCGCGCGGACAGCGCTCAAGCACCGCGAGT 7759
QY 603 GTCATGGGCGAGGGATCTCATATGACAGCGGCTCGATGTTCAACAGGGCGCTCGAGCT 662
Db 7758 GTCATGGGCGGGAAGATCAACATCGACTCGGCGACCTTTATGAACAAGGGCTCGAGGT 7699
QY 663 GATCGAGACGCGGAATCTTTCGCTTCGAGCGGACCGGATCGAGCGGTCTGTCATCC 722
Db 7698 ATTCGAGCGCATCTGGCTCTTCGACATGCGCGGACAGAAATCGGCTGCTGGTGCACCC 7639
QY 723 GCAATCCATCGTCCATGCGATGTTGGCTTCTTCGCGAGGGGCGCTGATGCCCATCTCGG 782
Db 7638 GCAGAGCATCATTCACTCGATGTTGAGTACATCGACGGCTGCTCATCGCCACGCTCGG 7579
QY 783 CCGCGCGGACATGCGCCACGCCATCGATTCGCGCTGAATGCGCGGCTCGCGCGAGGT 842
Db 7578 CGTGCCCGGACATGCGCGCTCCCATCGCTATGCGCTGCGCGGAGGCGCTCGGAAAC 7519
QY 843 GCCCGTCCCGGATCGACCTCGCACAGATTGCGAGCGCTCACTTCACAGAGCTCGACGA 902
Db 7518 CGGCATCGGCAAGCTCGATCTCACCAGGTGCGCCACACTCACTTCGAGGAGCGCAGACAT 7459
QY 903 GGAACGCTTTTCGCGCGCTTGAAGCTTTCGCGAGAGCGTATGCGCGCGCGCGCTGTGCGG 962
Db 7458 GGAGCGCTTCCCGCGCTTCGCGCTTCGAGCGCTTCGAGCGCTCAAGCGCGCGCAGACATATCC 7399
QY 963 CGCGGCTTCAACGCGGCCAAGAGATCGCGCTCGCATTTTCATCGCGGAGCGCATCGG 1022
Db 7398 TCGCGTACTGAACGCGGCCAACGAAATTCGCTGTGCGGCGGCTTCCTCTGCACAGAAATTCG 7339
QY 1023 GTTTCCTGACATGCGCGCGGTGTCGAGGAGAC 1055
Db 7338 CTTTACCGACATCGCGGAGCGGTGCGATAAAC 7306

RESULT 8
AL646064 201050 bp DNA linear BCT 07-DEC-2001
LOCUS Ralstonia solanacearum GMI1000 chromosome, complete sequence;
DEFINITION
```

segment 8/19.
 AL646064 AL646052
 VERSION AL646064.1 GI:17428340
 KEYWORDS
 SOURCE
 ORGANISM
 Ralstonia solanacearum.
 Ralstonia solanacearum
 Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 Ralstonia.
 1 (bases 1 to 201050)
 Salanoubat,M., Genin,S., Artiguenave,F., Gouzy,J., Mangenot,S.,
 Arlat,M., Billault,A., Brottier,P., Camus,J.C., Cattolico,L.,
 Chandler,M., Choisme,N., Claudel-Renard,C., Cunnac,S., Denange,N.,
 Gaspin,C., Lavie,M., Moisan,A., Robert,C., Saurin,W., Schlex,F.,
 Siguer,P., Thebaud,P., Whalen,M., Wincker,P., Levy,M.,
 Weisenbach,J. and Boucher,C.A.
 Genome sequence of the plant pathogen Ralstonia solanacearum
 Unpublished
 2 (bases 1 to 201050)
 Boucher,C.A.
 Direct Submission
 Submitted (05-DEC-2001) Genoscope and CNRS UMR-8030, 2 rue Gaston
 Cremieux, CP5706, 91057 Evry Cedex, France, Laboratoire de Biologie
 Moleculaire des Interactions Plantes-Microorganismes INRA-CNRS,
 BP27, 31326 Castanet-Tolosan Cedex, France, Fondation Jean
 Dausset-CRPH, 27 rue Juliette Dodu, 75010 Paris, France, LMGM CNRS
 118 Route de Narbonne, F 31062 Toulouse Cedex 4, Genoscope and INRA
 URGV, 2 rue Gaston Cremieux, CP5706, 91057 Evry Cedex, France,
 Laboratoire de Biometrie et Intelligence Artificielle INRA, BP27,
 F31326 Castanet-Tolosan Cedex, Laboratoire de Genetique Cellulaire
 INRA, BP27, F31326 Castanet-Tolosan Cedex
 Christian.Boucher@toulouse.inra.fr
 http://sequence.toulouse.inra.fr/R.solanacearum.html.
 Location/Qualifiers
 1..201050
 /organism="Ralstonia solanacearum"
 /strain="GMI1000"
 /db_xref="taxon:305"
 206..1447
 /gene="RscI326"
 /note="RS02848"
 206..1447
 /gene="RscI326"
 /EC_number="2.6.1.1"
 /function="small molecule metabolism; amino acid
 biosynthesis"
 Gene name confidence : probable
 /note="Product confidence : probable
 Gene name confidence : hypothetical
 predicted by Codon_usage
 predicted by Homology
 predicted by Framed"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="PROBABILE ASPARTATE AMINOTRANSFERASE PROTEIN"
 /protein_id="CAD15028.1"
 /db_xref="GI:17428341"
 /translation="MKTIQKAKLNVCYDIRGPVLEKAKQMEEGHQIIFKLNIGNLA
 PFGDPAPETIQDMIRNLPSAGYSDSGKIFAPKAAVMHYTQOQIKNVTLDIYLG
 GASLIALATNALLDAGDELLLPADYPLMTAMTSLSGGTPVHLCDESGNWPDLDD
 IRAKTPNKGIVLVNPNNSKYSRSGYRAGWMVSGDKRPADKYEGLNMLSSMRL
 RHTMASLVDLVITNLSKYSRSGYRAGWMVSGDKRPADKYEGLNMLSSMRL
 CANVPGWAIATAGYQSIINDLVAPGRMRQRDLAYELITAPGVSCVKKPAALYM
 PRLDPSPVYIPDDQTFIROLLEERVLLVQGTGNNHSPDHFRIVFLPHEDDLREAI
 GRIARFLERYRQRHGTGIRA"
 1538..2857
 /gene="RscI327"
 /note="RS02849"
 1538..2857
 /gene="RscI327"
 /EC_number="1.1.1.3"
 /function="small molecule metabolism; amino acid
 biosynthesis; threonine"
 /note="Product confidence : probable

```
/function="small molecule metabolism; biosynthesis of
cofactors, carriers; molybdopterin"
/notes="Product confidence : putative
Gene name confidence : putative
predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PUTATIVE MOLYBDOPTERIN BIOSYNTHESIS MOEA
PROTEIN"
/protein_id="CAD15032.1"
/db_xref="GI:17428345"
/translation="MGANDAPAVKPSLLTLEAALLLOLAAARGLTETVPTLHANGR
VLAQAVRSLGIPGADVMOECKPDPTGTWVIDHPEACEMIRAGEDITAGAEILPR
RTLGAQLGLAASVGCAGLVVRRVAVFFGGDELMPEKPLKPAIYNSRFLR
ALLENJCEIDTIGVFDTLQATRDRLREAAEADLIITSGVSGVEEDHDKPAVEAE
GRLENLQAIKPKGLAFGEVRRPQMGAGPAAFFILGPNPVSSVFTLLFVRPFI
RLGVDDVTPKRIPMRADFALPKGRNRLRINAGGGLDLFPNQSGVLTSTVMG
DGLIDNPNOPIAAGDVTAPLPFAGLV"
310. .6573
/gene="moaD"
/notes="RSC1331; RS02853"
6310. .6573
/gene="moaD"
/function="small molecule metabolism; biosynthesis of
cofactors, carriers; molybdopterin"
/notes="Product confidence : probable
Gene name confidence : probable
predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PROBABLE MOLYBDOPTERIN MPT CONVERTING FACTOR
(SUBUNIT 1) PROTEIN"
/protein_id="CAD15033.1"
/db_xref="GI:17428346"
/translation="MDMQIELRPFASVREQLGTSHAAAAPDPTVRTVGLRRLAARG
PANGELAEGRALRMALDHDVVTADAPLREGCEVAFPPVPTG"
6586. 7116
/gene="moaE"
/notes="RSC1332; RS02854"
6586. 7116
/gene="moaE"
/function="small molecule metabolism; biosynthesis of
cofactors, carriers; molybdopterin"
/notes="Product confidence : probable
Gene name confidence : probable
Query Match 24.7%; Score 292.2; DB 1; Length 201050;
Best Local Similarity 57.0%; Pred. No. 9,1e-25;
Matches 625; Conservative 0; Mismatches 448; Indels 24; Gaps 4;
```

QY 1 ATGCGCAGCTGTGATCTTTGGGCGCACCGGTCTCCATCGCGGAATCCACCTTCGACCTC 60
Db 97933 ATGATGCGCATTTACCGTTCTTGGCGCGCACCGGTCTCCATCGCGGACGCGCTG 97992

QY 61 GTCATGCGAAGGCGGCGCGGCGGTTCGCGACCGTCTGACCGCGGCGCGCAAC 120
Db 97993 GT-----GGCGCGCATTCGCGCGGTTCACCGGTGTTCGACTGACCGCAACACGCG 98046

QY 121 ATCGCGGACTGCGCGGAATGGCGGTGCGCTGAAGCGCGGAGCTTGGCGTACCGCGCAT 180
Db 98047 GTCGTAAGCTTGGCGCGGTGCGCGGTGTTCCGTCGCGGAGTGGCGGTGGTGGCGTCG 98106

QY 181 GAGACTGCTCGCGCGCTGCGGAGCGCTGGCGGG-----ACGGGCAACCGAGGTC 234
Db 98107 GCCACCGCGCGGAGGTGTCGCGGGACCGAGCTCGGCGCGGAGGCGCACCGGCGCATC 98166

QY 235 GCGGGCGGGCGAGGCCATGCGCGAGGCGCGCGCGCGCGCGGCTGACCATGTCG 294
Db 98167 CGCTTCGCGCGCGAGCGCTGGAAGAGCGCGCGCGCGCGCGGATGCGGATCGCGTATG 98226

QY 295 GCCATCGTGGCGCGCGCGCGCTGCTGCCGGGAATGCGGGGCTGAAGACGCGCGCGACG 354
Db 98227 GCGCCCATGCTGCGCGCGCGCGCGCTGCGCCGACGCTGCGCGCGCGTGGCGCGCGGAAG 98286

QY 355 C---TGGCGCTCGCAACAAGGAAGCTCTGTGAGCGCAGGCGCAACTCCCTCATCGCGGACG 411
Db 98287 CGTGTGCTGCGCAACAAGGAAGCGCTGTGTGATGTCCGCGCGCTCTTCATGAGCGCG 98346

QY 412 GCCCAGAGAACGCGCGCGCGACGATCTTCCGCTGACAGCGAGCAGCTCCGCGGTCTTTTCAG 471
Db 98347 GTGCGCGAGCATGCGCGCACCGCTGCTCCGATCGACAGCGCATACGCCATCTTCCAG 98406

QY 472 CGCTGCGCGCGGAGGACACGGCT-----GCCTGAGCGCGCTCATCATCGCGCG 522
Db 98407 TGCTTGCGCGAGCAGCGCTCCGAGCTTCGGGACGCTGCGCGCGCATCTGCTGACGCGG 98466

QY 523 TCGCGCGCGGTTCGCGGACTCGAGGCTCGAGGCGCTCGAGGCGCATCGCGCTGCACCGCTGCCGCGAG 582
Db 98467 TCGCGCGCGGTTCGCGGCGCGCTCGAGGCGCTCGAGGCGCTCGCGGCTGCGCGCGCATCG 98526

QY 583 GCGCAGGCG 642
Db 98527 GCCTGCG 98586

QY 643 TTCAACAAGCGCTCGAGCTGATCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 702
Db 98587 ATGAACAAGCGCGCTGAGGCTGATCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 98646

QY 703 ATCGAGCGCGCTCGCTCCATCGCGCAATCCATCGCTCCATCGCTGCTGGCGGTCTTTCGCGCGG 762
Db 98647 CTCGAGGTGCTGATCCACCGCGAGCGCTGATCCATTCGATGCTGCTTACGATGACGCT 98706

QY 763 GGCTGTGATGCCCATCTCG 822
Db 98707 TCGGTCTGCGCAACTTGGGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 98766

QY 823 TGGCGCGGTGCGCGGAGGTGCG 882
Db 98767 TACCCGAGCGCATGAGCG 98826

QY 883 ACCTTCAGAGCTGACGAGGAACGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 942
Db 98827 CGCTTGAAGCGCGCGATCTCGACCGCTTCCCGTGTCTGCGCTTTCGAGCGCGCATG 98886

QY 943 GCG 1002
Db 98887 CG 98946

QY 1003 TTCATCGCGCGCGCATCGCGGTTCGACATGCGCGCGGTGCTGCGAGGAGAGCGCTCGCG 1062
Db 98947 TTTTTCAGCGCGCATCCGCTTCCCGAGATGCGCGCGGTGCTGCGCGCGCGCGCGCGCG 99006

QY 1063 GCGGTTCGACCG 1079
Db 99007 CGCACG 99023

RESULT 9
AE012236/c
LOCUS
DEFINITION Xanthomonas campestris pv. campestris str. linear BCT 23-MAY-2002
of 460 of the complete genome.
ACCESSION AE012236 AE008922
VERSION AE012236.1 GI:21112424
KEYWORDS
SOURCE Xanthomonas campestris pv. campestris str. ATCC 33913.
ORGANISM Xanthomonas campestris pv. campestris str. ATCC 33913
Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
Xanthomonas.

REFERENCE
AUTHORS

- 1 (bases 1 to 10029)
da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R.,
Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida
Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,
Chambergio,F., Ciapina,L.P., Ciccarelli,R.M.B., Coutinho,L.L.,
Cursino-Santos,J.R., El-Dorri,H., Faria,J.B., Ferreira,A.J.S.,
Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C.,
Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite
Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A.,
Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J.,
Menck,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M.,
Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A.,
Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spinola,L.A.F.,
Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos
Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and
Kitajima,J.P.

TITLE

Comparison of the genomes of two Xanthomonas pathogens with

JOURNAL
MEDLINE
PUBMED

Nature 417 (6887), 459-463 (2002)

REFERENCE

22022145

AUTHORS

- 2 (bases 1 to 10029)
da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R.,
Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida
Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,
Chambergio,F., Ciapina,L.P., Ciccarelli,R.M.B., Coutinho,L.L.,
Cursino-Santos,J.R., El-Dorri,H., Faria,J.B., Ferreira,A.J.S.,
Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C.,
Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite
Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A.,
Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J.,
Menck,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M.,
Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A.,
Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spinola,L.A.F.,
Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos
Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and
Kitajima,J.P.

TITLE

Direct Submission

JOURNAL

Submitted (28-NOV-2001) Departamento de Bioquímica, Universidade de
Sao Paulo, Av. Prof. Lineu Prestes 748, Sao Paulo, SP 05508-900,
Brazil

FEATURES

source

Location/Qualifiers

1. .10029
/organism="Xanthomonas campestris pv. campestris str. ATCC
33913"
/strain="ATCC 33913"
/db_xref="ATCC:33913"
/note="taxon:190485"
/note="pathovar: campestris"
complement(72. .2528)
/gene="oma"
/note="XCC1365"
complement(72. .2528)
/gene="oma"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="outer membrane antigen"
/protein_id="AA040663.1"
/db_xref="GI:21112425"

gene

.....

CDS

/translation="MKLLHFRLLSSVAGTQOHPQSTPTGCDMTRLPTRFLLALALA
AGLSLPAALAAEPFVASDIRVDGLORIAAGTVTYLPVNRGTVDDGKVADSRAL
YRTGFEDVQDRGNILIVTKRPAINKLTITGKNDIKSEELKGLGDIGLTFEGGT
FDRSLDRVQTELFRQYNNRKNVEITPTVSPDRNRVDVAIAKEGKAARHVN
IGTEKFANEDILENKEHNSWASRRDDQYSKEMSGDLKLSNGLDRGYDFSV
DSTQVSIQPKRDMFVSAGVTEGQYKISEIKVQDTVLPOADLERLVPKAGDIFSR
ALLEYTSDAINTLSNGYAFKVNPIPTNREDTVAVNLQVVPGRVSRVILYK
NSRSTDEVLREMRQFENTWYQAAIDRSKIRLORLGYFFESVDVETPAVPGSNQVDV
VYNKVTISFVGLGFSQAGVTTISVQLSQNNPLGGGNRVAVEARSTPQQRVAFS
YTNPFPTDGGVSLGYNLSWRLEDYSDFNATQFNSTNGSAQVYVCGVPTENTVTSAMIG
VDSNQITTFPCTTPAIVDYIQAIQTDTFTKAVRTFEGWARDTRNDFMPTGRMYQRVG
LETTALPGSTVEYKLYNQISNVPFIIPALVNLTRFELGYGDSYGRKSLTKVSGDEM
RPVATLGPYFENYAGTSNVRGFDENLTGRSEAINGFNRGQPLGGSLKTVSGVEM
YFKLFDSPSARIASFDFGNVYSDVDAFKANELRASTGVALLWRAPVGPISIVAFP
LKEDNDEIERLQTFGCGQ"
complement(2525. .3871)
/gene="XCC1366"
complement(2525. .3871)
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA040664.1"
/db_xref="GI:21112426"
/translation="MGDFIGSWMMIVSLGVLVTFHERGFHWAVRRCGVKVLRFVSGF
GKPLMRDRHGTEFATAAIPLGYSVXMLDEREGEVHPAREQAFNRKTVQRTAIVA
AGPIANLLTCLHMLWAMFVIGKQDYSATVGRADGLAAAGLVPRGIRVIRIDGRSVSSW
SDASMQITTAAMDRDQVLTAAEEGNGSEHTLRLSQLPAGDFERRVATLAGIGQEF
LQPPVQVWVAGSAADGVLPKQDRIVADGQPIRSAGEVPAQLAQALGQTGGTMEVA
REDDRLALEIAPRAKSPGOMLGVGFATAAPAYDSROOYGAFAVPAARIRETGKMTA
DSLGMKMLTGOASVKNISGPTVIARAANAERGVDWFLYFLGLLSLALINLMP
IPLDGGHLLYLIELVKGSPISERAMIAGQYVGLAVLAGMLGAFYNDILGLVPR"
complement(3898. .5088)
/gene="dxr"
complement(3898. .5088)
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="1-deoxy-D-xylulose 5-phosphate reductoisomerase"
/protein_id="AA040665.1"
/db_xref="GI:21112427"
/translation="MSGSLRRVAVFGATSGISAGALDVIARHPRLRASVLSAGSKV
DALALCVLHRPAHVADAAALPALRDGLRAAGLTQAHAGDQALDAAADACDTY
VAATVGAAGLSSTLAAAGKRLILANKESLVLAGELTTRTAAAGAGIIPIDSHSA
IFQCLRSCDASGRVRRVLTATAGSGPFRGRORAOAEVTPAOAVAKWGMGPKISVDS
ATLANKGLEVIEAHLPLGCEQIDVLPQSLVHSLVFEVDGSLAGLGLPDMRTTL
AVGLAWPVERVESGGLDQLQGRIDFEAPDTEAPCLRLANDALRAGGTAPAILNAA
NEVAVSAFLQCKVGLFATPAIVLHTLTTLQRONADTLNALLFADAEARRTTERALAH
SLHA"
complement(5091. .5918)
/gene="cdsA"
/note="XCC1368"
complement(5091. .5918)
/gene="cdsA"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="phosphatidate cytidyltransferase"
/protein_id="AA040666.1"
/db_xref="GI:21112428"
/translation="MTQRTVIAALIMAPLAICAIVLLPTQWLAAALILFTGLWEWL
KLGGIDSLPRTVLLMLNLLMWASVGSWVLFQIALVGVGVWLVALLWLRFFEN
FGADHGNQARALKLAAGTLAVLPAAALVLLHANPKGNLWLLTALTWVWADSGAY
FAGRRFGKHLAPRISPNKTIIEGLLGVAGIACAFEGVLALUTLHAIVPQLVLAAY
AVLASVIGDLFESLLKRRHAGAGSGTVPVGHGVDLRIDGVLAAALPVFALGKELIFG"
complement(5915. .6691)
/gene="upps"
/note="XCC1369"
complement(5915. .6691)
/gene="upps"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="undecaprenyl pyrophosphate synthetase"
/protein_id="AA040667.1"
/db_xref="GI:21112429"

gene
/translation="MASTAMHSEPSAAVPRHLAIIIMDGNRWAORRRPRVIGHRAG
ARAVNRTIGDCLDKGVSALTFAFSENNWQDEVDALMKLFLHALDRGVEELORRG
VQVRIIGDRSRTAPLRDMAGAEIRITANTRILVLSIAASGYGRDIDIAAARALAVE
AAGRLOPEQIDALLASVALADPAPDLFIRTGDDTRISNELLWOLAYTELWFTETL
WPEFDAGVLQOALDDYAGRERREGLTSAQIAEKATEASSA"
complement(6694..7251)
/gene="frf"
/note="XCC1370"
complement(6694..7251)
/gene="frf"
/note="Identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="ribosome recycling factor"
/protein_id="AA040668.1"
/db_xref="GI:21112430"
/translation="MNLQIKDQAOITRMKSIDALRSLTIVTGRASPALDNIKVA
YGTDPILNOVASISVSEGRSLVIFDKMIDKEKAIYASDLGTPVVGTVIRNL
PPLTERRKELSKSVHGEDEKAVIRNIRRDANOQVKDILLKQVTEDEARGAEDDI
OKLTDKAIKDQDVVYKREQLMTV"
complement(7488..8210)
/gene="pyrH"
/note="XCC1371"
complement(7488..8210)
/gene="pyrH"
/note="Identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="uridylylate kinase"
/protein_id="AA040669.1"
/db_xref="GI:21112431"
/translation="MSELSYRILILKSLSEALMGDXYCIDPKVINRLAHEVIERQQA
GAQVALVIGGGFIRFRAIRHLEKRIATFAAGTGNPFTDSDGALRALEIGADILLK
ATKVDGVYDKPKKISDAVRVDSLEYDEVIMQGLEVMTAFAALARDSLPLRIGMS
EPGLVRLIHQAQIGTLVQGRS"
complement(8318..8983)
/gene="XCC1372"
complement(8318..8983)
/gene="XCC1372"
/note="Identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
Query Match 24.6%; Score 290.6; DB 1; Length 10029;
Best Local Similarity 57.2%; Pred. No. 3.2e-24;
Matches 612; Conservative 0; Mismatches 439; Indels 18; Gaps 4;
QY 2 TGGCAGCCTGTCGATCTTTGGGCGCACCGGCTCCATCGCGGAATCCACCTTCGACCTCG 61
Db 5072 TTGCTCGGGTCGCGGTGTTGCGGCGGACCGGCTCCATCGGTGCTCGGCACCTGGAGCTGA 5013
QY 62 TCATCGCGAAGGCGGCGCGGCGGCTTCGCGACCGCTCTGACCGCGGCGCGCAACA 121
Db 5012 TTGCGCGCA-----TCGGAGCGGCTGCGTCCAGCGTGTGTCGCGCGGCGCAAGG 4959
QY 122 TCGCGGCGACTGCGCGCAATGGCGCTGCGCTGAAGCGGAGCTTGCGCTACGCGCATG 181
Db 4958 TCGATCGCTGCTGCGGCTTGTGCGTCTGCACCGCCCGCCGCGGCTGATCGCGGACG 4899
QY 182 AGGACTGCTGCGCGGCTGCGGAGGCGCTG-----GCCGGAGCGGCACCGAGCTG 235
Db 4898 CCGCGTGTGATCGCGATTCGCGGATGTTGCGCGCGCGGCTGACCAACCGCGCGC 4839
QY 236 CGGCGGCGGCGGCGGCGGCGGCTGCTGCGGAGGCGCTG-----GCCGGAGCGGCACCGAGCTG 292
Db 4838 ATCGGCTGATCAGCGCGCTGGAGCGGCTGCGGCGGCGGCGGCTGCGGAGCGGCTG 4779
QY 293 CGGCCATCGTGGCGGCGGCGGCGGCTGCTGCGGGAATGCGGCGGCTGAAGACGCGCGCA 352
Db 4778 CGGCGATCGTGGCGGCGGCGGCGGCTGCTGCGGCGGCGGCGGCGGCGGCGCA 4719

QY 353 CGGTGCGGCTGCGCAACAAGAAAGCCTCGTACGCGAGGCGCAACTCTCTGATCGGACGG 412
Db 4718 GCCTGTGTTGGTGGCAACAAGAAATCCCTGCTGTGCGCGGCGAGCTGCTACCCGACCG 4659
QY 413 CCAGAGGAACGCGCGCCACGATCCTCGCGTGGACAGGAGCACTCCGCGCTTTTCAGG 472
Db 4658 CGGCTGCGGCGCGCGGAGATCATTTCCGATCGACAGCGACAGTGGCGATTTTCAGT 4599
QY 473 CGGTG---GCGGCGGAGGACACGCGCTGCTGAGCGCGCTCATCATCACGGCTCCGGCG 529
Db 4598 GCCTGCGCTCTGCGAGCGCAGCGGCGGCTGCGCGGCTGATCTCTGACGCGCTCCGGCG 4539
QY 530 GCGCGTTCGCGGCTGAGCGCTCGAGCGATCCGCGCTGACCGCTGCGGCGGCGGCGGAG 589
Db 4538 GCGCGTTCGCGGCGCGCGGCGGCTGCGCAATTGCGCGGCGGCGGCGGCGGCGGCGG 4479
QY 590 CCATCCCAACTCGTCCATGGCGCAGCGGATCTCCATCGACAGCGGCTCGATGTTCAACA 649
Db 4478 CGCACCCGAATGCTCGATGGGCGGCGGAGATCTCGGTGCGACTCGGCGCACCTGATGAACA 4419
QY 650 AGCGCTCGAGCTGATCGAGCGCGGGAATTTCTTCGGCTTCGAGCGCGGCGGCGGAGG 709
Db 4418 AGGCGCTGGAAGTATCGAGGCGCACCATCTATTGCGCTGCCCGGCGAGCAGATCGAG 4359
QY 710 CGGTGCTCATCGCGAATCATCGTCATCGATGCTGGCTTCTCGGAGGCGGCGGCTGA 769
Db 4358 TGTGCTGTCACCGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4299
QY 770 TGGCGCATCTCGGCGCGCGGCGGATCGCGCACCGCATCGGATTCGCGCTGAACTGCGCG 829
Db 4298 TGGCGCATCTCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4239
QY 830 GTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 889
Db 4238 AGCGCTCGAGTCCGCGCGTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4179
QY 890 AGAGCTGACGAGGAGCGCTTTCGCGCGCTGAGGCTTTCGCGGAGAGCTGATCGGCGCG 949
Db 4178 AAGTCCGAGACCGGAGGCGGCTTCCGCTGCTGCGCTGCGGCGGCGGCGGCGGCGG 4119
QY 950 GCGGCGCTGTCGCGGCGGCGGCTTCAACGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1009
Db 4118 GCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4059
QY 1010 CCGGCGCATCGGCTTTCGAGCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1058
Db 4058 AGGCGAAGGTGGGTTTCTTAGCCATCCCTGCGTGGTGGCAACACAGCT 4010
RESULT 10
AF282879
LOCUS
DEFINITION
Pseudomonas aeruginosa 1191 bp DNA linear BCT 15-SEP-2000
reductoisomerase gene; complete cds.
ACCESSION
AF282879
VERSION
AF282879.1 GI:9664862
KEYWORDS
Pseudomonas aeruginosa.
Pseudomonas aeruginosa
Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
Pseudomonas.
REFERENCE
1 (bases 1 to 1191)
AUTHORS
Altincicek,B., Hintz,M., Sanderbrand,S., Wiesner,J., Beck,E. and Jomaa,H.
TITLE
Tools for discovery of inhibitors of the 1-deoxy-D-xylulose
5-phosphate (DXP) synthase and DXP reductoisomerase: an approach
with enzymes from the pathogenic bacterium Pseudomonas aeruginosa
FEMS Microbiol. Lett. 190 (2), 329-333 (2000)
JOURNAL
MEDLINE
PUBMED
20487113
11034300
REFERENCE
2 (bases 1 to 1191)
AUTHORS
Altincicek,B., Jomaa,H. and Beck,E.
TITLE
Direct Submission

JOURNAL Submitted (27-JUN-2000) JLU-Giessen, Biochemisches Institut,
Friedrichstr. 24, Giessen 35392, Germany
FEATURES
Source Location/Qualifiers
1. .1191
/organism="Pseudomonas aeruginosa"
/db_xref="taxon:287"
CDS
1. .1191
/function="biosynthesis of isopentenyl diphosphate, IPP"
/note="Dxp reductoisomerase; Dxr"
/codon_start=1
/transl_table=11
/product="1-deoxy-D-xylulose 5-phosphate reductoisomerase"
/protein_id="AAF97241.1"
/db_xref="GI:9664863"
/translation="MSRPQRISVLGATGSGISLTLDVVORHPRDYAEFALTGFSRLAE
LEALCLRHVPVAVYVPEQAAALALGSLAAAGIRTVLFGEOALCEVASAPEDVMA
AIVGAGLPSTLAAYEAGKRVLLANKALVMSGALFMQAVRSGAVLPIDESHNAIF
QSLPRYADGLERVRRILLTASGPFRETLEQLASVPEQACAHNWSMKRISV
DSAMNKGLEIEACWLFDAQPSQVYVHPQSVIHSMDYVDVSGVIAQLGNPDMRT
PISYAMWPERIDSGVPLDMFAVGLDFORPEQRFCLRLASQAARETGGAPAMLN
AANEVAAFLERHFRISDIAVIEDVLNREAVTAVESLDQVLAADRARSVAGQWL
RHAG"

BASE COUNT 176 a 400 c 413 g 202 t

Query Match 24.2%; Score 286; DB 1; Length 1191;
Best Local Similarity 55.8%; Pred. No. 1.9e-23;
Matches 597; Conservative 0; Mismatches 455; Indels 18; Gaps 2;

QY 20 TTGGGCGACCGGCTCATCGGCGAATCACCTTCGACTCGTCATCGCGAAGGGCG--- 76
DB 29 TCGGCGCGACCGGCTGATCGGCTGAGCACCTCGACGCTGCCAGCGCTCATCCGGATC 88
QY 77 GGCCCGAGGGGTTCCGACCGTCTGACCGCGGCGGCGACATCCCGCGACTGCGCG 136
DB 89 GTTACGAGGCTTTCGCGCTGACTGGCTTCAGCGCGCTTGGCGCACTCGAGGCGCTGTGCC 148
QY 137 AATGCGCGTGGCTGAAGCGGCGAGCTTGCGCTCACCGCGCATGAGGACTGCGTCCCGG 196
DB 149 TCAGGCACCGCCCGCTTATGCGGTGTCGCGGAGCAGCGCGCGGCGATTGCGCTTCGAGG 208
QY 197 CGTGGCGGAGGGCTGGCGGCGACCGGACCGAGTTCGGGCGGCGCGAGCCATCG 256
DB 209 GCTCGCTCGCGCGGCGGGGTATCCGCCCGGGTGTGTTCGCGGAGCAGGCGTTCGCG 268
QY 257 CCGAGCGCGCGACCGCGCGGCGGCTGACCATGTTCGCGCATCTGCGCGCGCGCGCGGCG 316
DB 269 AAGTGGCAGCGCGCGGAGTGGACATGTTATGGCGGCCATCTGCGCGCGCGCGCGGCG 328
QY 317 TCCTGCGCGGAAATGCGGCGGCTGAAGCAGCGCGCGCATCGTGGCGCTCGCCCAACAGGAA 376
DB 329 TGCGCTGACCGCTTGGCGGCGCTGAGCGCGCGGCAAGCGCGTACTGCTGGCCCAACAGGAG 388
QY 377 GCTCTGCGGAGCGCACTCTCTATCGGAGCGCGCGGAGGAGAGCGCGCCACCATCC 436
DB 389 CGTGTGTGATGTCGCGGCGGCTGTTTCATGAGCGGCTGCAAGCGCGCGCGGCTGCTCC 448
QY 437 TGCGGCTGAGCAGCAGCACTCGCGGCTTTTCAGGCGCTGGCG-----G 481
DB 449 TGCGGATGACACGAGCAGCAAGCGCATCTTCAGTCTGCTGCGCGCGCAATTATGCCGATG 508
QY 482 GCAGGACAGCGGCTCGGTGAGCGGCTCATCATCAGCGGCTCCGCGGCGCGGCTTCGCGG 541
DB 509 GCCTGGAGCGGTCGCGGTCGCGCGGATCTCTTTCAGCGGCTCCGCGCGCGGCTTCGCGG 568
QY 542 ACTGGAGCTCGAGCGCATCCGCGCTTGACCGTGGCGGCGGCGAGGCGCCATCCCAACT 601
DB 569 AGACGCGCTGGAGCAACTCGCTTCGCTGACGCGGAGCGGCTTTCGCGCACCGCAACT 628
QY 602 GTTCCATGGCGGAGTCTCCATCAGCAGCGGCTTCGATGTTCAACAAGGCGCTCGAGC 661
DB 629 GGTCGATGGGCGTAGATTTCCGTGCTGCTCCGCCAGCATGATGAACAAGGGGCTCGAAC 698

QY 662 TGATCGAGAGCGCGCAATTCCTCGGTTTCGAGCGGACCGGATCGAGCGGTCGTCCATC 721
DB 689 TGATCGAGCGGCTGCTGGCTTCGAGCGGACCGGACGAGTTCGAGGTGATCCACC 748
QY 722 CGCAATCCATCGTCCATCGGATGGTGGCTTCTCGGACGCGGCGGCTGATGCCCATCTCG 781
DB 749 CGCAGAGCGTGAATCCACTCGATGGTGGACTACGTCAGCGTTCGGTATCGCCAGCTCG 808
QY 782 GCCCGCGCGACATCGCGCCACCGCATCGGATTCGCGTGAACCTGGCGGGTTCGCGCGGAGG 841
DB 809 GCAATCGGACATCGCGACCGCATTCCTATGCCATGGCTGGCGGAGGATCGATT 868
QY 842 TGCCCGTCCCGCGGATCGACCTCGCACAGATTCGAGCGCTCACCTTCCAGAACGCTGACG 901
DB 869 CCGGCGTTTCGCGCGTGGATATCTTCGCGCTCGCTCGCTGGATTTCACGCGCCGACG 928
QY 902 AGGAACGCTTTCGCGCGCTGAGGCTTCGCGGAGACGCTCATGCGGCGGCGGCGCTGTCGG 961
DB 929 AGCAGCGCTTCCCTCGCTGCGCTGCGAGCGGCGGAAACCGGCGGCGGAGCGGCC 988
QY 962 CGCCCGCTTCAACGCGCGCAAGGAGATCGCGTTCGATCATTTTCATCGCGCGGACGATCG 1021
DB 989 CGCCATGCTGAATGCGCGGACGAGTGGCGCTGGCGCGATTCTTCGAGCGGCATCC 1048
QY 1022 GGTTCCTGACATGGCGCGGCTGGTTCGAGGAGACGCTCGCGGCGGTTTCG 1071
DB 1049 GCTTCAGCGACATCGCGGTTATCATCGAGGACGCTGTGAACCGCGGAGCG 1098

RESULT 11
...AE004785/c
LOCUS
DEFINITION Pseudomonas aeruginosa PA01, section 346 of 529 of the complete genome
ACCESSION AE004785
VERSION AE004785.1
KEYWORDS GI:9949809
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
REFERENCE 1 (bases 1 to 14537)
AUTHORS Stover,C.K., Pham,X.-Q.T., Erwin,A.L., Mizoguchi,S.D., Warrenner,P., Hickey,M.J., Brinkman,F.S., Hufnagle,W.O., Kowalik,D.J., Lagrou,M., Garber,R.L., Goltry,L., Tolentino,E., Westbrook-Wadman,S., Yuan,Y., Brody,L.L., Coulter,S.N., Folger,K.R., Kas,A., Larbig,K., Lim,R., Smith,K., Spencer,D., Wong,G.K., Wu,Z., and Paulsen,I.T.
TITLE Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
JOURNAL Nature 406 (6799), 959-964 (2000)
MEDLINE 20437337
PUBMED 10984043
REFERENCE 2 (bases 1 to 14537)
AUTHORS Stover,C.K., Pham,X.-Q.T., Erwin,A.L., Mizoguchi,S.D., Warrenner,P., Hickey,M.J., Brinkman,F.S., Hufnagle,W.O., Kowalik,D.J., Lagrou,M., Garber,R.L., Goltry,L., Tolentino,E., Westbrook-Wadman,S., Yuan,Y., Brody,L.L., Coulter,S.N., Folger,K.R., Kas,A., Larbig,K., Lim,R., Smith,K., Spencer,D., Wong,G.K., Wu,Z., Paulsen,I.T., Reizer,J., Saier,M.H., Hancock,R.E.W., Lory,S. and Olson,M.V.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2000) Department of Medicine and Genetics, University of Washington Genome Center, University Of Washington, Box 352145, Seattle, WA 98195, USA

FEATURES
source Location/Qualifiers
1. .14537
/organism="Pseudomonas aeruginosa"
/strain="PA01"
/db_xref="taxon:287"
gene complement(66. .1418)
/gene="PA3649"
CDS complement(66. .1418)
/gene="PA3649"


```
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAG07037.1"
/db_xref="GI:9949810"
/translation="MSALTYIVGTTLVALGVLTFFHEFFHWARRCGVVKVLRFSVCGF
TPIVRHHRGTEFVVAIPLGGYVYKMLDERAEVPAHLLEQSFNRTVRQRIATVAA
GPTANFLLAIFWVVALGSOQVRVIGSVAPESLAQAQLEAGQELLAVDGEPTVG
WNGVNLQVLRRLGESGTELVROEKGSNVDSTHOVRLDGWLKAGEDNPDPITASLIGRPW
RPALPVLAELDPKPAQAAGLKLGRLOS IDGIAVDDMQVYVDSVRARPGORVOLKV
LRDGEVLDVALEAVRGECKARSGYMGAGVAGTEWPAEMLREVSYGPLEANGQALSRT
WTMSLTLDISIKMLGELSVRNLSGPTITIAKVAGASAGSGVGDFFLFLAYLSIGV
LNLPLIPVLDGGHLLPYLVEWNRPLSERVQAWGMQIGISLVVGVVMLLALVNDLSRL
"
complement(1444..2634)
/gene="dxx"
/notes="PA3650"
complement(1444..2634)
/gene="dxx"
/codon_start=1
/transl_table=11
/product="1-deoxy-d-xylulose 5-phosphate reductoisomerase"
/protein_id="AAG07038.1"
/db_xref="GI:9949811"
/translation="MSRPQISVLGATSGISGLSTLDVQVORHDDRYEAPALTGFSRLAE
LEALCUIHRHPVYAVVPEQAAATIALOGSLAAAGIRTVLFGEQALCEVASAPEVDMYMA
AIVGAAGLPSTLAAVEAGKRVLLANKLEALVMGALFMQVAKRSGAVLLPIDSEHNAIF
QSPRNYAGDLERGVGRIRILLTASGGPFRETPEQLASVTPEQACAHPNQSMGRKISV
DSAMNKGLELLEACWLFDAPQSOVEVTHPQSVIHSMDVYVDSVIAQLGNPDMRT
PISYANWPERIDSGVPLDMFAVGRDLDFQDEQREPCRLASQAETGGSPAPMLN
AANEVAFAFLERHIAFRSDIAVIEBVLNREAVTAVESLDQVLAADRARSVAGQWLT
RHAG"
complement(2631..3446)
/gene="cdsA"
/notes="PA3651"
complement(2631..3446)
/gene="cdsA"
/codon_start=1
/transl_table=11
/product="phosphatidate cytidyltransferase"
/protein_id="AAG07039.1"
/db_xref="GI:9949812"
/translation="MLKQRIITALLPITAGLFGFFLEGAFFALFICAVVSLGAWETA
RLAGYEQQGRVAYATVAVMLVALYHLPLQAGAVLLALVMTLATVLTLYTPESVG
YMGRRVRLGMGLILLPAWQGLVLLKQWPLANGLIAYVVLVWGADIGAYFSGKAFG
KRKLAPRVSPGKSWEGVYGGIAASLAIITLAVGLYRGWSLGAALLALIGAAVVFVSIV
GDITSMFKRQSGIKDSSNLLPCHGGVLDRLDLSLTAAIPVFAALLWAAGWGA"
complement(3440..4195)
/gene="upps"
/notes="PA3652"
complement(3440..4195)
/gene="upps"
/codon_start=1
/transl_table=11
/product="undecaprenyl pyrophosphate synthetase"
/protein_id="AAG07040.1"
/db_xref="GI:9949813"
/translation="MEKTRKDCVPRHVAIIMDGNNRWAKRLLPLGAVGHKAGVDVAVR
AVTECAAGAEVLTILFAFSSENKQRPADVSALMELFLVALRREVRLKLDENGIRLRI
IGDRTFPHPELOAMBEAETAAGTNTRELLQVAAVNGGOWDIIQAAQRLAREVOGGHL
AADDISALLQGLCVTGDQPLDLCINTGGEHRI SNFLLWQLAYAEIYSDLEFWPDK
HAAMRAALADFSSKRQRFRTGKTSQVEAEARHPC"
complement(4211..4768)
/gene="frr"
/notes="PA3653"
complement(4211..4768)
/gene="frr"
/codon_start=1
/transl_table=11
/product="ribosome recycling factor"
/protein_id="AAG07041.1"
/db_xref="GI:9949814"
/translation="MNEIKEAQERMGKLTLEALGHAFKIRTGRAHPSILDSVMVSY
YCADTPLROVANVTVEDSRTLALAVFDKSKIQAVKAIMTSDLGSLNPATAGTTIRVPM
PALTEETKRGYTKQARAEQARVSVNRIRRALAQKLDLQKEISEDEERRAGDOV
QKLTDFEIGEIEKALEAKEADLMAY"
complement(4771..5508)
/gene="pyrH"
/notes="PA3654"
complement(4771..5508)
/gene="pyrH"
/codon_start=1
/transl_table=11
/product="uridylylate kinase"
/protein_id="AAG07042.1"
/db_xref="GI:9949815"
/translation="MAQQLSARQPRYKRILLKLSGALMGSEEFIDPKVLDKMALEI
GQLVGIGVQVGLVIGGNLFRGAALSAGMDRVTDGDMGLMFLATVNLGLAMRDALERSN
IPALVMSAISMVGTVDHYDRRKAMRHVGGEVVFISAGTGNPFPTTDSACLRAI EID
ADVLLKATKVDGYTTADPFKDPNAEKFERLTDEYDLDRKLGVMDLTAICLCRDQNMPL
RVFNMKPGALLNIIVGGAESTLIEG"
complement(5707..6576)
/gene="tsf"
/notes="PA3655"
complement(5707..6576)
/gene="tsf"
/codon_start=1
/transl_table=11
/product="elongation factor Ts"
/protein_id="AAG07043.1"
/db_xref="GI:9949816"
/translation="MAEITAAVMYKELRERTGLGMMECKKALTAAGGDIKKAIDMRAA
GAIKAAKAGNTAAEGSIAVKIAADNKAAVIIEVNSOTDFLALQDDFKGFAESLEKA
FNEKLTDAAPLYAREEARLALVAKTGENVIRBLTRVGDVVGAYLGHRIQGVVNL
KGNPBLAKDIAHVAASNPQFLSASEVSEETAKEIFEFLALNADKIAKGPENIVEN
MYKGRISKFLAEASLVEQPFVKNPVKVGLAKQAKGAEIVSFVRYEVEGIEKAEVDF
AAEVAQAQAATQ"
complement(6707..7447)
/gene="rpsB"
/notes="PA3656"
complement(6707..7447)
/gene="rpsB"
/codon_start=1
/transl_table=11
/product="30S ribosomal protein S2"
/protein_id="AAG07044.1"
/db_xref="GI:9949817"
/translation="MSQVNRMDLKGAVHFGHGTWRVNPWKGKFIIGARNKIHIINLE
KITLPMEALTYFVERLAAGKNKILFVTKRSAGKIVREAAKCMFVVDHRLWGLMLT
NYKTIQSKTKRLDLETQSDGTFDKLTKEALMRSDLEKRLERSLGGIKDMGLPDA
LFVIDVDHERIAITEANKLGIPIGVVDTNSSPESGYDVIIPGNDDAIRAQVLYLNSWA
EAVIRKQGNATSADEFVEEAPAESAE"
7712..8497
/gene="map"
/notes="PA3657"
7712..8497
/gene="map"
/codon_start=1
/transl_table=11
/product="methionine aminopeptidase"
/protein_id="AAG07045.1"
/db_xref="GI:9949818"
/translation="MTVTIKTPDDIEKMRAGRLAAEVLMEIGEHKPGVYTTTELDR I
CHDYIVNEOKAIPAPLNYKGFPSICTSINHVVCHGIPNEKPLKEGDLNVDITVIKD
GYHDSYCKGLVKTEWADRLCQITQECNYKGISVVRPCAHLGDIGELIIOKHAENG
FSVVRKSGHGIGKVFHEEPQVILYGRAGTGIELKGMFTIETPMINOGRPETRLIGD
GWTAITDKRKLKSAQWEHTVLVTDAGYEIILFLRNDFTFPRTSAA"
8654..11366
/gene="glnD"
/notes="PA3658"
8654..11366
/gene="glnD"
/codon_start=1
/transl_table=11
/product="protein-PII uridylyltransferase"
/protein_id="AAG07046.1"
```



```
/db_xref="gi:9949819"
/translation="MPQVDFELFDRGQFAELALKSP IAAFFKKAIRQFREVLDNRFN
SGRIRRIEDRAWVDIQLQAAWQFDDADIALVAVGVYGRGLHPYSDVDLLI
LLDSEDFPREPTEGFTLLWDIGLEVGSVQOQAEARADLVITLMECRTI
GGPSLRQMLQVTSAAHWSPSEKFFLAKRHEQORRAKYNDEYNPNVKGSPGL
RDQITILMARQVGSNLHALVREGFLVESEKMLASQEFLLRVYLAHMLAGRAE
DRLLFDHORSIARLFQEDNDKLAVERFMOKYRYVMAISELNDLIQHFEFVILPC
EQVOIOPLSRFLRGDYIEVTHPNFKPTFALLIEIFVLMQHPETKYRADTIRL
LRDSRHLIDDFRIDIRNTSLFIELKSSQCIHRNLRMRYGILRVLPEFGHIGQ
MOHDLFTYTVDAHTLNLIKHLRNPMAEKYPLASKIIDRLPKPELIYIAGLYHD
IAKGRGDHSELGADAEFCQSHQPLWDLQVLSWLVQNLVYSTAQRKDLSDPOV
IFDPAQLVGDTHDLYLVLTADINATNPILMNSWRASLRLQYITETKRALRGLN
```

Query Match 24.2%; Score 286; DB 1; Length 14537;
Best Local Similarity 55.8%; Pred. No. 9.4e-24;
Matches 597; Conservative 0; Mismatches 455; Indels 18; Gaps 2;

QY	20	TTGGGGCCACCGCTCCATCGCGGAATCCACCTTCGACCTGCTCATGCGGAAGGCGG---76
DB	2606	TCGGCGCGACCGGCTCGATCGGCCTGAGCACCTCGGACGTCGTCCAGCGTCATCCCGCATC 2547
QY	77	GGCCCGGAGCGCTTCGCGACCGCTCTCACCGCGCGGCGCAACATCGGCGGACCTGGCGG 136
DB	2546	GTTAGAGACCTTCGCCCTGACTGCTTCAGCGCGCTGCGCGAATCGAGGCGCTGTGCC 2487
QY	137	AAATGGCGCGTTCGCTGAAGCGGAGCTTTCGCGTACCAGCGCATGAGGACTGCTTGGCCG 196
DB	2486	TCAGGACACCGCGCTATGCGGCTGCTGCGGAGCAGCGCGCGGATGCTTGGCTTGCAGG 2427
QY	197	CGCTGGCGGAGCGCTGCGCGGAGCGGACCGAGGTGCGGGCGGGGCGGAGGCCATCG 256
DB	2426	GCTCGCTGCGCGCGCGGCTATCCGACCCGCGGCTGCTGTCGCGGAGCAGCGGCTTGTGCG 2367
QY	257	CCGAGGCGCGCGCGCGCGGCGGCGGACTGGACCATGTCGGCCATGTCGGCGCGCGCGGCG 316
DB	2366	AGTGGCCAGCGCGCGCGGAGTGGACATGTAATGGCGGCATGTCGGCGCGCGCGGCG 2307
QY	317	TGCTGCGCGGAATGCGGCGCTGAAGCAGCGCGCGGCGGCTGCGGCTGCGCCAAAGGAAA 376
DB	2306	TGCCGTCGACCTGCGCGCGCGCTGAGCGCGCGGCGGCAAGCGCTACTGTCGCGCAAGGAGG 2247
QY	377	GCTGTGACGCGGCGCAACTCTTATGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 436
DB	2246	CGCTGTGATGTCGCGCGCGCTGTCATGCGGCGGCTCAAGCGGCGGCGGCTGCTCC 2187
QY	437	TGCGGTCGACGAGCAGCAGCTCCGCGGCTTTTCAGCGCGCTGCGG-----G 481
DB	2186	TGCCGATCGACAGCGAGCAGCGATCTTCCAGTCTGCTGCGCGCAATATGCGGATG 2127
QY	482	GCGAGACACGCGCTGCTGCGGCGGCTATCATCATCAGCGGCTGCGCGCGGCGGCTTCCGCG 541
DB	2126	GCTGTGAGCGGCTGCGCGGCTGCGCGGATCTCTTGACCGCTTCCGCGCGGCTTCCGCG 2067
QY	542	ACTGGAGCTTCAGCGCATCCGCGCTGACCGTGGCGGAGCGGCGGCGGCGGCGGCGGCT 601
DB	2066	AGACGCGGCTGAGCAACTCGCTTCCGCTGAGCGGCGGAGCGGCTTGTGCGCACCGCACT 2007
QY	602	GGTCTATGGCGGCGGATCTCCATCGACGAGCGCTTCATGTTCAACAGGCGCTCGAGC 661
DB	2006	GTCGATGGGCGTGAATTTCCGTCGACTCGCGCAGCATGATGACAGGCGGCTCGAC 1947
QY	662	TGATCGAGCGCGGAATTTCTTTCGCTGCGGCGGAGCGGATCGAGGCGGCTGCTCCATC 721
DB	1946	TGATCGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1887
QY	722	CCCAATCCATCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 781
DB	1886	CGCAGGCGTATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1827
QY	782	GCGCGCGCGGATGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 841
DB	1826	GCAATCCGAGCATGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1767

QY	842	TGCCCGTCCCGCGGATCGACCTCGCACAGATTTCGAGCGCTCACCTTCAGAGCGCTGACG 901
DB	1766	CGGCGCTTTCGCGCGCTGGATATGTTCCGCGCTCGGCTGGATTTCAGAGCGCGCGGACG 1707
QY	902	AGGAACGCTTTCGCGCGCTGAGCGCTTCGCGGAGACGCTCATGCGGCGCGCGCGCTGCGG 961
DB	1706	ACGACGCTTCCCTTCGCTGCGCTGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1647
QY	962	GGCGCGCTTCAACGCGGCGGAGGAGATCGCGCTCGATCATTTTCATCGCGGCGGCGGCGG 1021
DB	1646	CGGCGCATGCTGAATGCGCGGAGAGGTCGCGTGGCGCGCATTTTCGAGGCGGCGGCGG 1587
QY	1022	GCTTTCGACATGCGCGGCTGCTGAGGAGCGCTCGCGGCGGCTTCG 1071
DB	1586	GCTTCAGCGACATCGCGGTTATCATCGAGGAGCTGCTGAACCGCGGAGGCG 1537

RESULT 12
SC5H4
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS

Streptomyces coelicolor cosmid 5H4.
AL355913.2 AL645882
AL355913.2 GI:20520782
1-deoxy-D-xylulose 5-phosphate reductoisomerase; acyl CoA dehydrogenase; aldehyde dehydrogenase; ATP/GTP binding protein; beta-galactosidase; cytidine deaminase; gcpE; glycosyl transferase; integral membrane protein; kinase; lipoprotein; metalloproteinase; prolyl tRNA synthetase; regulatory protein; secreted sugar hydrolase; sugar hydrolase; transcriptional regulator; transcriptional termination/antitermination factor; translation initiation factor IF-2; transport system integral membrane protein; two-component system response regulator; two-component system sensor kinase.
Streptomyces coelicolor A3(2).
Streptomyces coelicolor A3(2).
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
1 (bases 1 to 39744)
Redenbach, M., Kieser, H.M., Denapait, D., Eichner, A., Cullum, J., Kinashi, H. and Hopwood, D.A.
A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome
Mol. Microbiol. 21 (1), 77-96 (1996)
97000351
8843436
2 (bases 1 to 39744)
Murphy, L. and Harris, D.
Unpublished
3 (bases 1 to 39744)
Cerdano, A.M., Parkhill, J., Barrell, B.G. and Rajandream, M.A.
Submitted (11-MAY-2000). Streptomyces coelicolor sequencing project, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof. David A. Hopwood, [3] John Innes Centre, Norwich Research Park, Colney, Norwich, Norfolk NR4 7UH, UK
On May 9, 2002 this sequence version replaced gi:7801248.
Notes:
Streptomyces coelicolor sequencing at The Sanger Centre is funded by the BBSRC and Beowulf Genomics
Details of S. coelicolor sequencing at the Sanger Centre are available on the World Wide Web.
(URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are numbered using the following system eg SC7B7.01c. SC (S. coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary strand).
The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous. The length in codons is given for each CDS.
Usually the highest scoring match found by fasta -o is given for CDS which show significant similarity to other CDS in the database. The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon. Gene

prediction is based on positional base preference in codons using a specially developed Hidden Markov Model (Krohn et al., Nucleic Acids Research, 22(22):4768-4778(1994)) and the FramePlot program of Bibb et al., Gene 30:157-66(1984) as implemented at <http://www.nih.go.jp/~jun/cgi-bin/frameplot.pl>. CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, ttg or (att)) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions. Cosmid 5H4 lies between and overlaps cosmids 8B7 and 9F2 on the AseI-B genomic restriction fragment.

FEATURES

Location/Qualifiers
1. .39744
/organism="Streptomyces coelicolor A3(2)"
/strain="A3(2)"
/db_xref="taxon:100226"
/clone="cosmid 5H4"
complement(2..136)
/notes="nominal overlap with Streptomyces coelicolor cosmid SCB87"

misc_feature
689..836
/notes="nominal overlap with cosmid St5H4 from 1 to 148"

misc_feature
1219..2877
/gene="SC5H4.02"
/notes="SCO5678"
1219..2877
/gene="SC5H4.02"
/notes="SC5H4.02, possible regulatory protein, len: 552 aa; similar to TR:Q9ZBP0 (EMBL:AL034492) Streptomycetes coelicolor putative regulatory protein SC6C5.05c, 569 aa; fasta scores: opt: 310 z-score: 323.3 E(): 1.5e-10; 30.0% identity in 587 aa overlap"
/codon_start=1
/transl_table=11
/product="putative regulatory protein"
/protein_id="CAB91114.1"
/db_xref="GI:7801250"
/db_xref="SPTREMBL:Q9KYT7"
/translation="MPTPLASLVHHSALKLTVRAGEDRLDVPVRAHVSSELADPPVPM EGELLITALKLDADPEAMRYVRKLAGVVGGLGFAVGVNDVPEALVDAAQOE GLPLLEVPRTPELTSKAVSAIAADQAVTAGFAOQRELTTRTLTGPEGLAAL AQVDGWAALPDAGAVASAPPEWAGRAARLTADYORLRERPARPASSVVGAGNTGH SENADRVLSLGTSSRRPRAALAVGTAAALGTAERYAVHSATLTLTTERSLSLHEA GLKIDAAVLRMLLAGEPDHARTVAGDLYGLLDADPPRVIVAESGAARNRGPTAAHTAA GTGGDPPGLTGLSLAAEASAAAGEAVLAVDPGERLVVLPATDGGAAVACVHE ASALEAARPAEPLAGGDEESLVVGSAPSGPIAASAAKQAQQAALSVARRRGVCVE HEHVAAGSVLPILADDVAFADGLRLRDHDATGCGDLVASVRAWLSRHGQWDAAD. ADLGVHRHTLVYRMRRVEELGRLSDPDVRLMLALKATAPE"
3021..4466
/gene="SC5H4.03"
/notes="SCO5679"
3021..4466
/gene="SC5H4.03"
/notes="SC5H4.03, probable aldehyde dehydrogenase, len: 481 aa, similar to TR:Q66573 (EMBL:AF000677) Aquifex aeolicus aldehyde dehydrogenase AldH1, 476 aa; fasta scores: opt: 1129 z-score: 1260.2 E(): 0; 39.9% identity in 471 aa overlap and to SW:GAPN_STRMU (EMBL:L38521) Streptococcus mutans NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.9) GapN, 475 aa; fasta scores: opt: 893 z-score: 997.5 E(): 0; 35.4% identity in 477 aa overlap. Contains Pfam match to entry PF00171 aldedh, Aldehyde dehydrogenase family"
/codon_start=1
/transl_table=11
/product="putative aldehyde dehydrogenase"
/protein_id="CAB91115.1"

/db_xref="GI:7801251"
/db_xref="SPTREMBL:Q9KYT6"
/translation="MTSTHAFWLAGRAAGESGFDVTSFMDGTTTGVTSLSPTDAQVEE AVAAAVFAVRDEPAATPAHVRAAALDHVGRLEARTETRIARLISAENGKPKVMARGEVG RAVSVFRAEARENRNGEQAQRLDTDAGQGQRLATRRPKGVVIGIAPFNPPLNLC AHKIAIPAAGAAALILKIPAPATPLSLGILGELAEITLDPAGSILPVPKMPALVQ DELPLVISTGSETSVGYAINDSVPRKHCTLELGGNGAAVVLADWADDEDIDRAATRIA TFSNYGGQSCISVGYAINDAAVYDRLRLRIVAAVQAQVTGDPNDADTDYGPVLSSEA ABEVWEAREAVDAGAKLITGGKDGASVAPTVLTLPADTTLAHEEVFGPVLVSQVRV TGPAEFAAVANDSKYGLQAGVTFHDLQAAFAHRALEVGVGWIGIDIPSYRADQMPYGG AKOSGVRGSGVKFAMDDYYERVVLITGLAL"
3048..4454
/gene="SC5H4.03"
/notes="Pfam match to entry PF00171 aldedh, Aldehyde dehydrogenase family, score 357.30, E-value 1.6e-103"
complement(4476..4889)
/gene="SC5H4.04c"
/note="SCO5680"
complement(4476..4889)
/gene="SC5H4.04c"
/notes="SC5H4.04c, probable cytidine deaminase, len: 137 aa; similar to TR:Q53367 (EMBL:AL021841) Mycobacterium tuberculosis cytidine deaminase Cdd, 133 aa; fasta scores: opt 185 z-score: 244.6 E(): 3.6e-06; 37.4% identity in 115 aa overlap and to SW:CDD_MOUSE (EMBL:AA388666) Mus musculus cytidine deaminase (EC 3.5.4.5) Cdd, 146 aa; fasta scores: opt: 173 z-score: 229.1 E(): 2.6e-05; 29.2% identity in 113 aa overlap. Contains Pfam match to entry PF00383 dCMP_cyt_deam, Cytidine and deoxycytidylate deaminase zinc-binding region and match to Prosite entry PS00903 Cytidine and deoxycytidylate deaminases zinc-binding region signature"
/codon_start=1
/transl_table=11
/product="putative cytidine deaminase"
/protein_id="CAB91116.1"
/db_xref="GI:7801252"
/db_xref="SPTREMBL:Q9KYT5"
/translation="MTSOTNPVDHELIEAAAHVARTRCRGNHTMAAAARARDGRIVT AVNAVHTGTGCAELVVVGAAGGVLELTIVAVGDRGVVPPCGRCRQVLLDYFP GLEVIIVGEGDRVAPVPTDLLPETVYVADHQLDAG"
complement(4575..4820)
/gene="SC5H4.04c"
/notes="Pfam match to entry PF00383 dCMP_cyt_deam, Cytidine and deoxycytidylate deaminase zinc-binding region, score 10.30, E-value 0.027"
complement(4599..4724)
/gene="SC5H4.04c"
/note="PS00903 Cytidine and deoxycytidylate deaminases zinc-binding region signature"
complement(4915..6090)
/gene="SC5H4.05c"
/note="SCO5681"
complement(4915..6090)
/gene="SC5H4.05c"
/notes="SC5H4.05c, possible glycosyl transferase, len: 391 aa; similar to TR:Q9X7Z1 (EMBL:AL049497) Streptomyces coelicolor putative glycosyl transferase SC6G10.05c, 412 aa; fasta scores: opt: 1364 z-score: 1557.1 E(): 0; 56.4% identity in 385 aa overlap. Contains Pfam match to entry PF00534 Glycos_transf_1, Glycosyl transferases group 1"
/codon_start=1
/transl_table=11
/product="putative glycosyl transferase"
/protein_id="CAB91117.1"
/db_xref="GI:7801253"
/db_xref="SPTREMBL:Q9KYT4"
/translation="MSDSRARTLIVTNDPFRQGGIETFTVRELADRPDPDEVVLFS PTAAPGSGESFPYVVRHPVTRLLPTPGATAHAAAVVARYGCDRVWFGAAPTALMAG RLRTAGIGTAVASTVGHVWVWARTPGRRALLRRIGSOVDTLTGLERTRAPTEALAA AGTTRALVPGVDVTGTFHPAADGTVRVTRVYGLGHRPVLCAARLVPRKQDQTLRALP WYRRAPVDAVLLLVGDPHAGLRLALREGVWDSVHFAGGHPHALLPAFYAADAPA MPCRTKRGLEGLGIVYLEAASGLPVAAGSGAPDAVREGTGHVVDGKSVANT

Query Match	23.6%	Score 278.6	DB 1	Length 39744
Best Local Similarity	56.1%	Pred. No. 4.7e-23		
Matches 602	Conservative 0	Mismatches 444	Indels 27	Gaps 3
Qy 9	CCTGTCGATCTTTGGGGCCACCGCTCATCGGGAATCCACCTTCGACCTCGTCATCGG 68			
Db 24771	CGTGGTGTCTCGGTCCACCGGTGATCGGACCACGAGCCATCGACCTTCGTGCTCGC 24830			
Qy 69	GAAGGGGGGGCCCGAGGGCTTCCGACCGTCGTCTGTACCGCGCGGGCGCAACATCCGGCG 128			
Db 24831	CAACCGGACCGCTTCGCGGTTCACGGCCCTGTCCGCCAAGCGGCGCGGTGCGCCCTC-- 24888			
Qy 129	ACTGGCCGAATGGCGGTGCTGTGAAGCGGAGCTTCCGCTCAACCGCGCATGAGAGATG 188			
Db 24889	-CTCGCGAGCAGCGTACCGGCTGAAGCGCGGACCGTCGCGCTCGCGCGCGAGGACGT 24947			
Qy 189	CCTCGCCGCTGCGCGAGGGCGT-----GGCGGGAGCGGCACGA 230			
Db 24948	CGTCGGCGCTTCGGGAGGCGCTACCGCCAGTACGGCACGGGCGAGCGCTCCCCGA 25007			
Qy 231	GGTCGGGGGGGGCGAGGCGCATCGCGAGCGCGCCGACCGCGCGCGCGCTGACCAT 290			
Db 25008	GATCCTCGCGGACCGGAGGCGCACCCAGCTCGCGCCCTCCGACTGCCACACCGTCT 25067			
Qy 291	GTGCGGCATCTGTGGCGCGCGGGCTCGTGCCCGGAATGCGGGCGCTGAAGACGCGCG 350			
Db 25068	CAAGGGCATACCGGCTCATCGGACTCGCGCCACGCTCGCGCGCTCGGAGCGGGCGG 25127			
Qy 351	CACGCTGCGCTGCCACAGGAAGGCTCGTGACGCGAGGGCAACTCCTGATCGGGAC 410			
Db 25128	CACCTTCGCTGGCCACAAGGAGTGCCTCATGTCGGCGCGCGCTGGTCAAGGCGCT 25187			
Qy 411	GGCCAGGAGAACGGCGGCGACGACTCTGCCGTTGGACAGCAGCACTCCGCGGTCTTTCA 470			
Db 25188	GGCCAAAGCG-----GGCCAGATCATCCCGTTCGACTCCGAGCAGCGCCCTTTCCA 25241			
Qy 471	GGCGTGGCGGGGAGGACAGCGGCTGGTTCGAGGGCTCATATCATCGGGCTCGCGGG 530			
Db 25242	GGCGTGGCGCGCGGACCGCGGCGAGCTGCGGAAGCTGTGCTGCTACCGCTCGCGGG 25301			
Qy 531	GCGCTTCGCGACATGAGCCTCGAGCGCATCGCGGCTGACCGTGGCGGAGGCGCAGC 590			
Db 25302	CCGCTTCGGGGCGGTACAGGGAGAGCTGGGCGCGCTACCGTTCGAGGACGCGCTCGC 25361			
Qy 591	CCATCCCAACTGTGTCATGGCCAGCGGACTCTCCATCCACAGCGGCTCGATGTTCAACAA 650			
Db 25362	CCACCCCACTGGCCCATGGCGCGGTGATCAGATCACTCCGCGGACGCTCGTCAACA 25421			
Qy 651	GGCGCTGAGTGATTCGAGACGCGGAATCTTCGGCTTCGAGCGGACCGGATCGAGGC 710			
Db 25422	GGGCTGGAGTGATTCGAGGCCCCACCTGCTCTACGACATTCCTTCGACCGCATTTGAGT 25481			
Qy 711	GGTGTCCATCCGCAATCCATGTCATGCTGATGGTGGGCTTTCGACAGGGGGCGCTGAT 770			
Db 25482	GGTGTGCACCGCGAGTGTATGTCCACTCGATGGTTGAGTACAGGACGAGTCAACGCT 25541			
Qy 771	GGCCCATCTCGGCCCGCGCATGCGCCACGCGCATCGATTTCGGCTGCACTGCGCGGG 830			
Db 25542	GGCCACGCCACGCCCCCGACATGGGGGGGCCCATTCGCCCTGGGCTTCGGCTGGCCGA 25601			
Qy 831	TCGGGGGAGGTGCCGCTCGCGCGGATCGAGCTCGCACAGATTGGAGGCTCACCTTCCA 890			
Db 25602	ACGGTCCCAGCGCGCCCCCGCTTCGACTGGAGCAAGGCTCGACCTGGGAGTTCTT 25661			
Qy 891	GAAAGCTGACGAGAAAGCTTTTCGGGCCCTGAGGCTTCGCGGACGACGTATGGGCGCG 950			
Db 25662	CCCGCTGCACAACAGAGGCTTTCCTCCGCTCGACTGGCCAGGACGCTGGACAGCTCGC 25721			
Qy 951	CGGCTTCGGGGCGCGCTTCAAGCGGGCCAAAGAGATCGCGCTCGATCATTTTCATCGC 1010			
Db 25722	GGGACGGGCCCGCGGTTTCAATGCCCGCCCAATGAGGAGTGCCTCGAGGGGTTCCGCTC 25781			

QY 1011 CGGACGCATCGGGTTCTTGACATCCGCCGTGTCAGAGAGACTCAGCGG 1063
 | | | | ||| | | | |
Db 25782 CCGCGCGTCCGGTTCTCGGATCATGAGACCCTCACGGGGTGGTCGAGG 25834
 ||| | | | | ||| | | | |

```

RESULT 13
LOCUS      AE011773/c
DEFINITION Xanthomonas axonopodis pv. citri str. 306, linear BCT 29-MAY-2002 of
            the complete genome.
ACCESSION  AE011773
VERSION     AE011773.1 GI:21107580
KEYWORDS
SOURCE
ORGANISM    Xanthomonas axonopodis pv. citri str. 306.
Xanthomonas axonopodis pv. citri str. 306
Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group; xanthomonas.

REFERENCE
AUTHORS     da Silva,A.C.R., Ferro,M.I.T., Formighieri,E.F., Franco,M.C., Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C., Camargo,L.E.A., Camarotte,G., Cannavan,F., Cardoso,J., Chambergro,F., Ciapina,L.P., Ciccarelli,R.M.B., Coutinho,L.L., Kursino-Santos,J.R., El-Dorrry,H., Faria,J.J.B., Ferreira,A.J.S., Ferreira,R.C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A., Madeira,C.F.M., Miyaki,C.Y., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J., Menck,C.A.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M., Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A., Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spinola,L.A.F., Takita,M.A., Tamara,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and Kitajima,J.P.
```

TITLE	Comparison of the genomes of two Xanthomonas pathogens with differing host specificities
JOURNAL MEDLINE PUBMED	Nature 417 (6887), 459-463 (2002)
REFERENCE AUTHORS	22022145 12024217 2 (bases 1 to 10242) da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R., Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C., Camargo,L.E.A., Camarotte,G., Cannavan,F., Cardoso,J., Chambergro,F., Ciapina,L.P., Ciccarelli,R.M.B., Coutinho,L.L., Coursino-Santos,J.R., El-Dorrry,H., Faria,J.J.B., Ferreira,A.J.S., Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C., Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A., Madeira,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M., Menck,C.F.M., Miyaki,C.Y., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A., Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A., Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spinola,L.A.F., Takita,M.A., Tamara,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and Kitajima,J.P. Direct Submission Submitted (28-NOV-2001) Departamento de Bioquímica, Universidade de Sao Paulo, Av. Prof. Lineu Prestes 748, Sao Paulo, SP 05508-900, Brazil

```

FEATURES             source
Location/Qualifiers
1..10242
 /organism="Xanthomonas axonopodis pv. citri str. 306"
 /strain="306"
 /db_xref="taxon:190486"
 /note="pathovar: citri".
 .95..286
 /gene="xAcI1412"
 95..286
 /gene="xAcI1412"
 /note="putative; ORF located using Glimmer/Genemark"
 /codon_start=1
 /transl_table=11
```

gene

CDS


```
QY 137 AATGCGCGTGGCTGAAGCGGAGCTTTCGCGTCAACCGCGATGAGGACTGCTGCGCG 196
  || || || || || || || || || || || || || || || || || || || || ||
Db 5231 CGTGTGCGCGACCCACCGCGCTGGCGATGCGCTGATTCCGATCGAGCTGTACCGCG 5172
  || || || || || || || || || || || || || || || || || || || || ||
QY 197 CGCTGGCGAGCGCTGGCGGCGGACCGAGGTCGCGCGGCGGCGAGGCCATCG 256
  || || || || || || || || || || || || || || || || || || || || ||
Db 5171 CGCTGGCGATGCGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 5112
  || || || || || || || || || || || || || || || || || || || || ||
QY 257 CGAGGCGCGC-----CGACCGCGCGCGGCGGACTGACCATGTCGCGGCGGCGG 307
  || || || || || || || || || || || || || || || || || || || || ||
Db 5111 CGCTGACCGCACTGGCGGACGAGTGTGCGACACCGTGTGGCGCGGATCGTGGTG 5052
  || || || || || || || || || || || || || || || || || || || || ||
QY 308 CGCGGCGCTGCTGCCCGGAATTCGGGCGCTGAAGACAGCGGCGCACGTCGCGCTCGCCA 367
  || || || || || || || || || || || || || || || || || || || || ||
Db 5051 CGCGGCGGCTGCGCTACGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4992
  || || || || || || || || || || || || || || || || || || || || ||
QY 368 ACAAGGAAGCGCTGTCGACGAGCGCACTCTCTGATGCGGAGCGGCGGCGGAGACGGCG 427
  || || || || || || || || || || || || || || || || || || || || ||
Db 4991 ACAAGGAATCCCTGGTCTGGCGGCGGAGCTGCTGACCGCGACCGCGCGCGCGCGG 4932
  || || || || || || || || || || || || || || || || || || || || ||
QY 428 CCACGATCTGCGGTGGAGAGCGACACTCCGCGGTCTTTCAGGCGCTG---GGGCGG 484
  || || || || || || || || || || || || || || || || || || || || ||
Db 4931 CTGAATCATCCGATCGACAGCGAACAACAGCGCCACTCTTCCAGTCCGCTCTTGGC 4872
  || || || || || || || || || || || || || || || || || || || || ||
QY 485 AGACACGCGCTGCTGAGCGCGTCATCATCAGCGGCTGCGGCGGCGGCTTCCGCGACT 544
  || || || || || || || || || || || || || || || || || || || || ||
Db 4871 AGCGACGCGCGGCTGACCGCGGCTCATCTGACTGCTTCCGCTGCGGCGGCTTCCG 4812
  || || || || || || || || || || || || || || || || || || || || ||
QY 545 GGAGCTCTAGCGCATCCGCGCTGACCGTGGCGAGCGGCGGCGGCGGCGGCGGCGG 604
  || || || || || || || || || || || || || || || || || || || || ||
Db 4811 GGACCGCGCGGCTGGCGGCGGCTGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4752
  || || || || || || || || || || || || || || || || || || || || ||
QY 605 CCATGGCGCAGCGGATCTCCATCGACAGCGCTCTGATGTTTCAACAGGCGCTCGAGCTGA 664
  || || || || || || || || || || || || || || || || || || || || ||
Db 4751 CGATGGCTCGAAGATCTCGGTGATTCGCGACCTGATGATGACAGGCGCTTGGAGTCA 4692
  || || || || || || || || || || || || || || || || || || || || ||
QY 665 TCGAGACGCGGAATTTCTTGGCTTGGAGCGGACCGGATGAGGCGGCTGCTCCATCCG 724
  || || || || || || || || || || || || || || || || || || || || ||
Db 4691 TCGAGCGCATCACCTGTTTGGACTGCGCGGCGGAGCAGATCGATGCTGTGTCATCCG 4632
  || || || || || || || || || || || || || || || || || || || || ||
QY 725 AATCCATCTCCATGATGTTGGGCTTCTGCGAGGCGGCGCTGATGCGGCGATCTCGCG 784
  || || || || || || || || || || || || || || || || || || || || ||
Db 4631 AGAGCTGTGATTCCTTCCGTTGGGATTCGTCGAGCGGCTCCACCTTGCCTGAGCTGG 4572
  || || || || || || || || || || || || || || || || || || || || ||
QY 785 CGCGCGCATGCGCGCACGCGATTCGCGCTGAACCTGAGCGGCTGCGGCGGCGGAGGTGC 844
  || || || || || || || || || || || || || || || || || || || || ||
Db 4571 TCGCGGACATGCGACACCTTGGCGGTGCGGCTGCGGTGCGGCGGAGCGGTGCGATCG 4512
  || || || || || || || || || || || || || || || || || || || || ||
QY 845 CGTGGCGCGGATGACCTTCGCACAGATTCGAGGCGCTACCTTCCAGAGCGCTGAGAGG 904
  || || || || || || || || || || || || || || || || || || || || ||
Db 4511 GCGTCGCGCGGTTGGATCTGCTCAGCCAGGCGCGGTGATTCGAGGCGGCGGACACCG 4452
  || || || || || || || || || || || || || || || || || || || || ||
QY 905 AACGCTTTCGCGCTGAGGCTTGGCGGAGAGTATGCGCGGCGGCGGCGGCTGTCGGGG 964
  || || || || || || || || || || || || || || || || || || || || ||
Db 4451 CGCATTCCTCATGCTGCGCGCTGGCGTGGAGCGGTACGCGCGGCGGCGGCGGCGGCGG 4392
  || || || || || || || || || || || || || || || || || || || || ||
QY 965 CGCGCTTCAACCGCGGCAAGAGATCGCGCTCGATCATTTTCATCGCGGAGCGATCGGGT 1024
  || || || || || || || || || || || || || || || || || || || || ||
Db 4391 CGATCTGAACGCGGCGCATGAATGGCTGTTTTCAGCGTCTTTCAGGCGGAGGTGGGT 4332
  || || || || || || || || || || || || || || || || || || || || ||
QY 1025 TTCTGACATGCGCGGCTGGTTCGAGGAGACGCT 1058
  || || || || || || || || || || || || || || || || || || || || ||
Db 4331 TCCTAGCATCCCTGGCTTGGTTCGAACACACGCT 4298
  || || || || || || || || || || || || || || || || || || || || ||
```

RESULT 14

AB049187

LOCUS

DEFINITION

Streptomyces griseolosporeus dxr gene for 1-deoxy-D-xylulose

5-phosphate reductoisomerase, complete cds.

AB049187

ACCESSION

AB049187.1

VERSION

GI:13434972

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

Streptomyces griseolosporeus DNA.

Kitasatospora griseola

Bacteria; Actinobacteriales; Actinobacteridae; Actinomycetales;

Streptomyces; Streptomycetaceae; Kitasatospora.

1 (sites)

Hamano,Y., Daiiri,T., Yamamoto,M., Kuzuyama,T., Itoh,N. and Seto,H.

growth phase dependent expression of the mevalonate pathway in a

terpenoid antibiotic-producing Streptomyces strain

Unpublished

2 (bases 1 to 1161)

Daiiri,T.

Direct Submission

Submitted (22-SEP-2000) Tohru Daiiri, Toyama Prefectural University,

Biotechnology Research Center; 5180 Kurokawa, Kosugi-machi,

Imizu-gun, Toyama 939-0398, Japan (E-mail:daiiri@pu-toyama.ac.jp,

Tel:81-766-56-7500(ex.561), Fax:81-766-56-2498)

Location/Qualifiers

1..1161

/organism="Kitasatospora griseola"

/db_xref="taxon:2064"

1..1161

/gene="dxr"

1..1161

/gene="dxr"

/codon_start=1

/evidence=experimental

/transl_table=11

/product="1-deoxy-D-xylulose 5-phosphate reductoisomerase"

/protein_id="BAB39759.1"

/db_xref="GI:13434973"

/translation="MVIIGTSGISQIDVVLNRPKVVVALSAAGAVELLAQQA

VALGVHTVADPAEAEAAAGPGGAGRLPRVLGAPDAATELAEEHSLVNGIT

GSIGLAPTALALRAGRLVLANKEGLVGGPVKAVAGQGVIPVDSHAALFQALAG

GARAEVRLVVTASGPPRNTRLOLAATPADALAHPTWAMPVVTINSATLVNKL

EVEIAHLIDVDFRIEYVHPVSHMVEVDGSTMQAQSPDPMRPIALGLGWPD

RVPDAPGCDMTKATWEEFPLDNEAFVLAELREVLTGGTAPAVFNAANECDVAF

LKGALPFTGIVDTVAKVVAEHGTPOSGTSLATVEDVLHAEFNAARARELAAG"

BASE COUNT 151 a 399 c 449 g 162 t

ORIGIN

Query Match 20.9%; Score 247.2; DB 1; Length 1161;

Best Local Similarity 56.9%; Pred. No. 4.1e-19;

Matches 474; Conservative 0; Mismatches 353; Indels 6; Gaps 1;

QY 231 GGTGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 290

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 234 GGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 293

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 291 GTCGCGCATCGTGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 350

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 294 GAAGCGCATCGCGGTTTCGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 353

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 351 CAGCGTGGCGCTCGCCACAAGAAAGCCTCTGTCAGCGGCGGCGGCGGCGGCGGCGG 410

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 354 GGTGCTGGTGTGTCGGAACAAGAGTCGCTGATCTGTCGGCGGTCTGCTGAAGCGGT 413

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 411 GCGCCAGGAGAACGCGCGCCACGATCTGCGCGTGGAGACGAGCAGCTCCGCGGTCTTCA 470

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 414 GCGCGAGCC-----CGGCGAGATCGTGGCGGTGGACTCCGAGCAGCGCGGTTC 467

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 471 GCGCGTGGCGGCGGAGACACGCGCTGCTGAGCGCGGTATCATCATCAGCGGCTCCGCGG 530

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 468 GCGCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 527

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 531 GCGGTTCCCGGACTGAGCGCTCGAGCGCATCCGCGCTGCACCGTGGCGGCGGCGGCGG 590

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 528 CCGGTTCCCGCAACCGCACCGCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 587

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 591 CCATCCCAACTGCTCATGGCGGCGGCGGATCTCATTCAGCAGCGGCTCGATGTTCAACAA 650

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 588 GCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 647

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||


```
INERQFQGHFDPDLPVPCVLLIEAMAQACGTLAILSEGGKRENEFFFFAGIDEARFRK
QVDPDQDLVFEVVELLTSRRGIGKFNAAKVDGQVAVEAIMCAKRVV"
/translation=(4550..5596)
/gene="NMB0180"
CDS
complement(4550..5596)
/gene="NMB0180"
note="similar to PID:1718487 percent identity: 100.00;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="UDP-3-O-(3-hydroxymyristoyl)-glucosamine
N-acetyltransferase"
/protein_id="AAF40637.1"
/db_xref="GI:7225399"
/translation="MIPATYTLTSQITARLGLGWERGDISVTAVRPLADAAHEISFLA
HPKYAEVHDSAGAVIVSAKAADGFEGRNLIVADDPYLYFAKVARLFPSPVVKARGGI
NPTAVPEGATVPSTCEIGANVYIGANTLGEGRILANAVVQHDCKLGDSEVLHPNA
VYVYCTILGRREVHSGAVIGADGFLAADDNFKIPQTCGAVTLGDDVEIGSNTNID
RGMSDITVGNKIDNQVQIGHSCKIGSHVIAAKTIGISGVITGSCYCIIGGGVGV
GHIEDATTTIGGTSVTHSITGSKHLGIFPMSTHKEWARNNAVYIHRUSEMNRKLK
TLEQLSDAGQDSK"
complement(5629..6129)
/gene="NMB0181"
CDS
complement(5629..6129)
/gene="NMB0181"
note="similar to GB:M34854 SP:P31519 PID:155479 percent
identity: 46.88; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="outer membrane protein OmpH, putative"
/protein_id="AAF40638.1"
/db_xref="GI:7225400"
/translation="MTRLTAFRAALIGLCITAGAHADTFQKIGFINTERIYLESKOA
RKIOKTDSEFSAQDELQKREGLDLROLAEGKLNRNKAQAEKRWGLVAFRK
KQAFEDYLNRRNEFASLQANRVIVKAKQEGDVLQNVYIYNTQIDVDSVI
KEMNAR"
complement(6195..8588)
/gene="NMB0182"
CDS
complement(6195..8588)
/gene="NMB0182"
note="similar to GP:2460281 percent identity: 99.87;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="outer membrane protein Omp85"
/protein_id="AAF40639.1"
/db_xref="GI:7225401"
/translation="MKLKQIASALMLGISPLALADFTIQDIRVEGLQRTPEPSTFVNY
LPVKVDYNDTHGSAIKSLYATGFFDDVRVETADGQLLTIVIERPTIGSLNITGAK
MLQNDATIKLESFGLAOSQYFNQATLNQAVAGLKEEYLGRLKNIQITPKVTKLARN
RVDIDITIDEKSAKIIDIEFEGNOVYSDRLKMQMSLTGEGIWTWLTFRSNQFNEQKF
AODMEKVTDFYONNGYDFRILDTDIOTNEDKTKOTIKITVHEGGRFPWGVSTEGDT
NEVPAEAELEKLTWKPKGWYERQOMTAVLGEIONRMSGAVAYSEISVQPLPNAETKT
VDFVLHTEPGRKIYVNEIHTGNKTRDEVVRRELROMESAPYDTSKLQSKERVEILL
FQVFNQDFVLAPLQDPKVDLNSLTERSTGSLDSAGWQDVTGLVNSAGVSQDNLF
GTGSAALRASRSTYLLNSGSLDFDPTADGVSGLDYGVKAFDPKRAKSTSIKQYKT
ITAGAKIMSVPVTEYDRVNFGLVAEHLTVYTNKAPHYADFKFYKKTGDTGDSFKF
GWLKGTWGRNKTDSALWPTRGVLTGVNAEIALPGSKLOYYSATNHTOWFFPLSKT
TFLMGGEVLAGGYGRTKEIPFENYFEGGLSGVRGESLTLPKPYVDEIGKISYK
GNKANVASAELFPMPGAKDARTVRLSLFADAGSVWDGKTYDDNSSASTGVRQNIY
AGNTHKSTFTNELRYISAGAVTWLSPLGPMKFSYAYPLKKKPEDEIQRFPQLGTFP"
complement(8645..9985)
/gene="NMB0183"
CDS
complement(8645..9985)
/gene="NMB0183"
note="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF40640.1"
/db_xref="GI:7225402"
```

```
/translation="MHTLLAFIFAILILVLSHERGHIIVARLCGVKVVYRVSFVGFKPFF
FTRRQDTEWCLAPIPLGQYKVMVDREGEVSEADLYAFDKQHPKARIAIVAGPLT
NLALAVLYLGLSFQGYTELRYVGTVPDPTIARAGFQSGDKIQSVANGTPIVADMGSA
QTEIVLNLEAGVAVGTASGATVRIIDAAGTPEAGKIAKNOCYIGLMPFKITTTVA
GVEKGSPEAKGALPKGRLTAADGKPIASQEWANLTROSPGKITLNYBRAGOTHT
ADIRPDVTEQSDHTLILGRVGLRPQPDRAWDQAIRRSYRPSVYVRAFSGMEKTVSWT
TLKFFGKLISGNASVSHISGLPITTIADIAGSAELGLOSTLEFLALVLSGLVNLPLV
PVLDDGHLIVFYTAEMIRGPIGIERVQIIGLRFGLALMLMMAVAEFNDVTRLLG"
complement(10020..11204)
/gene="NMB0184"
CDS
complement(10020..11204)
/gene="NMB0184"
Query Match 20.2%; Score 238.2; DB 1; Length 13538;
Best Local Similarity 53.8%; Pred. No. 2.1e-18;
Matches 598; Conservative 0; Mismatches 483; Indels 30; Gaps 4;
QY 9 CTTGTGATCTTTGGGCGCACCGCTCCATCGCGGCAATCCACTTCGACCTGTCATCGG 68
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11190 CCTGACCATATTAGCAGTACCGCGCAGCATAGGCGAAAGCAGCTGAGACGTTCTCCC- 11132
QY 69 GAAGGCGGGCGCGAGGCGTTCGCGACCGCTCGCTGACCGCGGCGCAACATCCGGCG 128
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11131 -----GCCACCCCGCAAAATATCCGCGTATTCGCTGGCAGGCATAACAGGTCGAGAA 11077
QY 129 ACTGCGCGCAAAATGGCGCTGCGCTGAAGCGGAGCTTGCGCTACCGCGCATGAGGACTG 188
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11076 ATTGGCGGCTCAATGTCAAACGCTCCACCCGAATATGCGCTGTCGCGGATCGGGAACA 11017
...QY 189 CTTGCGCGCGCTGCGCGAGGCGCGT-----GCCGGGAGCGGACGAGGTGCGGGCGG 242
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11016 CCGCGCGCGCTTGAAGCGCTTGAAGCGGACGCGGCGGACTCAAGCTTTTACACGG 10957
QY 243 GCGCGAGGCGCATCGCGAGGCGCGC---GACCGCGCGCGGCGGACTGAGCACCATGTCGCCCAT 299
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10956 CCGCAGGCGCATGTTGTTGACGTTGCTCTCCACAGNAAGTCACGGGTGTCATGTGCGCCAT 10897
QY 300 CTTGGCGCGCGCGGCGCTGTCGCCGGAATCGCGCGCTGAAAGCAGCGGCGGCGCACGCTGCG 359
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10896 CCGTGGCGCGGTGGGCGCTGCTTCCGCGTCCGCGCGGCGCAAAAGGCAAAACCATTTA 10837
QY 360 GCTCGCCCAACAGAAAGCCTCGTGACGCGGCGCAACTCTTGATGCGGACGCCCGCAGA 419
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10836 TCTGCAAAAGAAAGACGCTGGTGTTCGCGCGCGGCTTTATGGAACCGCGCGCTGC 10777
QY 420 GAACGCGCGCACGATCTCGCGGTGGACAGCAGCAGCTCCGCGCTTTTCAGCGCGCTGCG 479
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10776 AAACGCGCGCGCAGTCTGCTCGCGTGCACAGCGAACAACACGCGCTTTCCAGTTTGGC 10717
QY 480 GGGCGAGGACACGG-----CTGGCGTGCAGCGCGCTCATCATCAGCGCGTC 524
Db 10716 GCGCGATTACGCGCGCGCTCTGAACGCAACACGCGCATCGCTTATCTGACCGCTTC 10657
QY 525 CCGCGCGCGCTTCCGCGACTGAGCGCTCGAGCGCATCCGCGCTCGACCGTGGCGCGAGC 584
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10656 CCGCGCGCGCTTCTGACCGCGCATTTAAACACGTTTCGACCGCGATTACGCCGCCCAAGC 10597
QY 585 GCAGGCGCCATCCCACTGCTGCTCCATGGCGCAGCGGATCTCCATCGACAGCGCGCTCGATGTT 644
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10596 GGTCAAAACACCCCAATTGGCGTATGGGACGCAAAATCTCCGTCGATTCGCCACCATGAT 10537
QY 645 CAACAAGCGCTCGAGCTGATCGAGACGCGGCAATTTCTTGGCTTCGAGCGCGGACCGGAT 704
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10536 GAACAAAGTTTGGAGCTGATTGAAGCGCATTTGGCTGTTTCAACTGTCCCGCGCGCAAACT 10477
QY 705 CGAGGCGGTCTCCATCCGCAATCCATCTCCATCGGATGGTGGCTTCTGCGACCGGGG 764
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10476 CAAAGTCGTCATCCATCCGCAATCCGTCATACAGCATGGTGGCTGCGCGACCGGCTC 10417
QY 765 CTTGATGGCGCCATCTCGGCGCGCGGACATGCGCCACGCGCATCGGATTCGCGTGAACATG 824
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10416 CGTGTGGCGCAACTGGGCAATCCCGATATGCAAGCGCCCATCGCTTATTTGGGTTT 10357
```


Result No.	Score	Query		DB	ID	Description
		Match	Length			
c	1	214.6	18.2	4403765	4	US-09-103-840A-2
	2	126.8	10.7	1197	4	US-09-491-362-6
	3	126.8	10.7	1197	4	US-09-874-562-6
	4	125.2	10.6	1732	4	US-09-449-335-5
	5	124.4	10.5	1434	4	US-09-480-921B-28
	6	123.6	10.5	1732	4	US-09-449-335-1
	7	122	10.3	1434	4	US-09-480-921B-9
	8	99.2	8.4	1759	4	US-09-491-362-1
c	9	99.2	8.4	1759	4	US-09-874-562-1
	10	74.8	6.3	8438	1	US-07-945-283-1
	11	72.2	6.1	2809	3	US-09-000-016-1
	12	72.2	6.1	2809	4	US-09-514-340-1
	13	71.6	6.1	5173	1	US-08-242-677-1
	14	69.4	5.9	20235	1	US-07-642-734C-3
	15	69.4	5.9	20235	3	US-08-439-009A-3
	16	66.6	5.6	2539	3	US-09-000-016-3
	17	66.6	5.6	2539	4	US-09-514-340-3
	18	66.6	5.6	11220	4	US-09-105-537-32
	19	66.6	5.6	36778	4	US-09-105-537-5
	20	66.6	5.6	38506	3	US-09-320-878-19
	21	65.6	5.5	1288	1	US-08-440-856A-9
	22	64.6	5.5	44377	2	US-08-804-227C-7
	23	64.6	5.5	44377	2	US-08-804-198-1
24	64.4	5.4	1187	1	US-08-440-856A-2	
25	63.2	5.3	1182	4	US-09-385-028-19	
26	63.2	5.3	11604	4	US-09-385-028-13	
27	63.2	5.3	15079	4	US-09-385-028-1	

NAME/KEY: CDS
LOCATION: (1)...(1197)
US-09-874-562-6

Query Match
Best Local Similarity 10.7%; Score 126.8; DB 4; Length 1197;
Matches 375; Conservative 0; Mismatches 382; Indels 9; Gaps 1;

Qy 293 CGGCCATCGTGGGCGCGCGGCTGCTGCGCGGAATGCGGCGCTGAAGCAGCGCCGCA 352
Db 299 CCGGAATAGTAGTGTGCGGCGACTAAAGCCTACGGTGTGCAATGAAGCAGGAAGG 358
Qy 353 CGTGGCGCTCGCAACAGAAAGCCCTGCTGAGCGGAGGCACTCTGTATGCGAGGG 412
Db 359 ACATGCTCTTGCACAAAGAGACATTAATCGAGGTGGTCTTCTGCTTCCGCTTG 418
Qy 413 CCCAGGAGACGGCGCCACGATCTGCGGTGGAGCAGGACCTCCGCGTCTTTCAGG 472
Db 419 CCAACAACATAAATGAAGATTTCTCGCGGAGATTGAGACATCTGCCATATTTCA 478
Qy 473 CGTGGGCGGAGGACACGCGCTCGCTGAGCGCGTCAATCATCACGCGCTCCGCGGGC 532
Db 479 GTATTCAAGGTTTGCTGAAGCGCTCTGCGCAAGATAATCTTGACTGCATCTGTGG 538
Qy 533 GCTTCCGCGACTGGAGCCTCGAGCGCATCCGCGCTGACCGTGGCGGAGCGGCGCC 592
Db 539 CTTTGTAGGATTTGGCTGCGGAGGCTGCGCAAGGTTAAAGGAGTAAAGTAGCGGAT 598
Qy 593 ATCCCACTGTGTCATGGCGCGAGGATCTCCATCGACGAGCGCTCGATGTTCAACA 652
Db 599 ATCCAACTGGAACATGGGAAGAAATCACTGTGGACTCTGCTACGCTTTTCAACA 658
Qy 653 CGCTCGAGCTGATCGAGACGCGAAATTTCTTGGGCTTCGAGCGCGGACGCGATCG 712
Db 659 GTCCTGAGGTCAATGAAGCGCATTAATTTGTTGGAGCTGAGTATGACGATATAG 718
Qy 713 TCGTCCATCCCAATCCATCTGTCATGCGATGGTGGGCTTCTGGAGCGGCGCTGATG 772
Db 719 TCATTCTCCGCAAGTATACATATCCATGATGTAAGTAAACACAGGATTCATCTG 778
Qy 773 CCCATCTCGGCGCGCGCGACATGCCACGCGATCGGATTCGCGTGAACCTGCGCGG 832
Db 779 CTCATTTGGTGGCTGATGCGTTTACCGATCTTACACCATGTCATGGCCGATA 838
Qy 833 GCGGC-----GAGGTGCGCGTGGCGGATGCGCTGACAGATTCGAGCGCTCA 883
Db 839 GAGTTCCTTGTCTGAAGTAACTTGCGCAAGACTTGACCTTTGCAAGCTCGGTTCA 898
Qy 884 CTTCCAGAGCCTGACAGGACGCTTTCCGCGCTGAGGCTTCGCGGAGACGTCATGG 943
Db 899 CTTTCAAGAAACAGACAAATGTGAATACCCATCCATGGATCTTGTCTTATGCTG 958
Qy 944 CGGCGCGCGCTGTCGCGCGCGCGCTTCAACGCGCGCGCAAGAGATCGCGCTCG 1003
Db 959 GAGCTGGAGGCACATGACTGGAGTCTCAGCGCGCGCAAGAGAAAGCTGTTGA 1018
Qy 1004 TCATCGCGGAGCATCGGTTTCTGGACATGGCGGCGGTGGTGA 1049
Db 1019 TCATTGATGAAGATAAGTATTTGGATATCTTCAAGGTTGTGA 1064

RESULT 4

US-09-449-335-5
Sequence 5, Application US/09449335
Patent No. 6303365
GENERAL INFORMATION:
APPLICANT: Busch, Marco
APPLICANT: Hain, Rudiger
APPLICANT: Martin, William
APPLICANT: Tietjen, Klaus
APPLICANT: Kloti, Andreas
TITLE OF INVENTION: Method of determining the activity of
1-deoxy-D-xylulose-5-phosphate reductoisomerase and

TITLE OF INVENTION: 1-deoxy-D-xylulose-5-phosphate synthase
FILE REFERENCE: 202005
CURRENT APPLICATION NUMBER: US/09/449,335
CURRENT FILING DATE: 1999-11-24
EARLIER APPLICATION NUMBER: DE 199 35 967.9
EARLIER FILING DATE: 1999-07-30
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 1732
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: CDS
LOCATION: (18)..(1448)
OTHER INFORMATION: new coding sequence
US-09-449-335-5

Query Match 10.6%; Score 125.2; DB 4; Length 1732;
Best Local Similarity 48.8%; Pred. No. 3.9e-15;
Matches 374; Conservative 0; Mismatches 383; Indels 9; Gaps 1;

Qy 293 CGGCCATCGTGGGCGCGCGGCTGCTGCGCGGAATGCGGCGCTGAAGCAGCGCCGCA 352
Db 550 CCGGAATAGTAGTGTGCGGCGACTAAAGCCTACGGTGTGCTGCAATGAAGCAGGA 609
Qy 353 CGTGGCGCTCGCAACAGAAAGCCCTGCTGACGGCAGGCACTCTGTATGCGAGGG 412
Db 610 ACATGCTCTTGCACAAAGAGACATTAATCGAGGTGGTCTTCTGCTTCCGCTTG 669
Qy 413 CCCAGGAGACGGCGCCACGATCTGCGGTGGAGCAGGACCTCCGCGTCTTTCAGG 472
Db 670 CCAACAACATAAATGAAGATTTCTCGCGGAGATTGAGACATTTCTGCCATATTT 729
Qy 473 CGCTGGCGGCGAGGACACGCGCTCGCTCGAGCGCTCATCATCGGCGTCCGCGGGC 532
Db 730 GTATTCAAGTGTGCTGAGCGGCTCTGCGCAAGATAATCTTGACTGCATCTGGT 789
Qy 533 GCTTCCGCGACTGGAGCCTCGAGCGCATCCGCGCTGACCGTGGCGGAGGCGAGGCC 592
Db 790 CTTTGTAGGATTTGGCTGTCGAAAGAGTAAAGTAGCGGATGCTGTTGAAGC 849
Qy 593 ATCCCACTGTGTCATGGCGCGAGGATCTCCATCGACGCGCTCGATGTTCAACA 652
Db 850 ATCCAACTGGAACATGGGAAGAAATCACTGTGGACTCTGCTAGCTTTTCAACA 909
Qy 653 CGCTCGAGCTGATCGAGACGCGGAATTTCTCGGCTTCGAGCGGACCGGATCGAG 712
Db 910 GTCCTGAGGTCAATGAAGCGCATTAATTTCTTGGAGCTGAGTATGACGATATAG 969
Qy 713 TCGTCCATCCGCAATCCATGCTGTCATGCGATGGTGGGCTTCTGCGAGCGGCGCTG 772
Db 970 TCATTCTCCGCAAGTATCATACATTCATGATTCATGATTTGAACACAGGATTCAT 1029
Qy 773 CCCATCTCGGCGCGCGCGCGCGCTTCAACGCGCGCGCAAGAGATCGCGCTCGAG 832
Db 1030 CTCATTTGGTGGCTGATATGCTGTTTACCGATTTCTTACACCATGTCATGGCC 1089
Qy 833 GCGGC-----GAGGTGCGCGCTGCGCGCGGATCGACCTCGACAGATTTGGAG 883
Db 1090 GAGTTCCTTGTCTGAAGTAACTTGCGCAAGACTTGACCTTTTCAACATCGGT 1149
Qy 884 CTTCCAGAGCCTGACGAGGAACGCTTTCGCGCGCTGAGGCTTGGCGGAGACGTCAT 943
Db 1150 CTTTCAAGAAACAGACAAATGTGAATACCCATCCATGGATCTTGTCTATGCT 1209
Qy 944 CGGCGCGCGCTGTCGCGCGCGCGCTTCAACGCGCGCGCAAGAGATCGCGCTCG 1003
Db 1210 GAGCTGGAGGCACATGAGTGGAGTTCACGCGCGCGCAAGTGAAGCTGTTGA 1269
Qy 1004 TCATCGCGGAGCATCGGTTTCTGGACATGGCGCGCGGTGGTGA 1049
Db 1270 TCATTGATGAAGATAAGCTATTTGGATATCTTCAAGGTTGTGA 1315

RESULT 5
US-09-480-921B-28
; Sequence 28, Application US/09480921B
; Patent No. 6387637
; GENERAL INFORMATION:
; APPLICANT: Levin, Joshua Z.
; APPLICANT: Budziszewski, Gregory J.
; APPLICANT: Potter, Sharon L.
; APPLICANT: Wegrich, Lynette M.
; TITLE OF INVENTION: Herbicide Target Genes and Methods
; FILE REFERENCE: PB/5-30780A
; CURRENT APPLICATION NUMBER: US/09/480,921B
; CURRENT FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 1434
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1434)
; OTHER INFORMATION: encodes SEQ ID NO:29
US-09-480-921B-28

Query Match 10.5%; Score 124.4; DB 4; Length 1434;
Best Local Similarity 48.6%; Pred. No. 5.4e-15;
Matches 372; Conservative 2; Mismatches 383; Indels 9; Gaps 1;

QY 293 CGGCCATCTGCGGGCGCGGCTCTGCGCGGAATGCGGGCGCTGAAGCAGCGCGCA 352
Db 533 CCGGAATAGTAGTTGTGCGGAGCTAGCGGTGTGCAATTTGAAGCAGGAAGG 592
QY 353 CGCTGGCGCTCGCAACAAGGAAGCCTCTGACGGCAGCACTCTCTATCGGACGG 412
Db 593 ACATTGCTCTTGCAACAAGAGACATTAATCGCAGGTGTCTTTCGTCTTCCGCTG 652
QY 413 CCCAGGAGAACGGCCACGACTCTGCGGTGGACAGCAGCACTCTCGCGTCTTTCAGG 472
Db 653 CCAACAACATAATGTAAGATTTCTCCGCGAGATTCAGAACATTCGCCATATTTCACT 712
QY 473 CGCTGGCGGCGGAGCAGCGCTGCGCGGTGGACAGCAGCACTCTCGCGTCTTTCAGG 532
Db 713 GTATTCAAGTTTGGCTTGAAGGGCGCTCTGCGCAAGATAATCTTGACTGCATCTGGTGGAG 772
QY 533 CGTTCCGCGACTGGAGCTCGAGCGCATCGCGCTGACCGTGGCGGAGCGCAGGCC 592
Db 773 CTTTGGGATTTGGCTGTCGAAGAGCTAAGGAAGTTAAGTAGCGGATCGCTTGAAGC 832
QY 593 ATCCCAACTGGTCCATGGCCAGCGGATCTCCATCGACAGCGCTCTCGATGTTCAACAAGG 652
Db 833 ATCCCAACTGGAACATGGGAAGAAAATCACTGTGACCTCTGTACGCTTTTCAACAAGG 892
QY 653 CGCTCGAGCTGATCGAGACCGCGAATCTTCGCTTCGAGCGGACCGGATCGAGCGG 712
Db 893 GTCTTGAGGTCATGTAAGCCCATTTATTTGTTGGAGCTGAGTATGACGATATAGAGATTG 952
QY 713 TCGTCATCCGCAATCCATCGTCCGATGTTGGCTTCTCGCAGCGGGGCTCTATGG 772
Db 953 TCATTATCCKAAAGTATCATATCATCTTCATGTTGAACACAGGATTCATCTGTGCTTG 1012
QY 773 CCATCTCGGCCCGCGGACATCGCCACGCCATCGGATTCGGGCTGAACTGGCGGGGTC 832
Db 1013 CTCAATTTGGGTTGGCTGATATCGGTTTACCGATTCTCTACACCATGTCATGSCCGATA 1072
QY 833 GCGG-----CGAGGTCCCGTCCCGGATCGACCTCGGCACAGATTCGAGCGCTCA 893
Db 1073 GAGTTCCTTGTCTGAAGTAACCTTGGCCWAGACTTGACCTTTGCAAACTCGGTTTCATGA 1132
QY 884 CTTTCCAGAGCTGACGAGGAACGCTTTCCGGCCCTGAGGCTTTCGCGAGAGCTCATGG 943

Db 1133 CTTTCAAGAAACCAGACAATGTGAATACCCATCCATCCATGATCTTGTCTTATGCTCTGGAC 1192
QY 944 CGGCGCGCGCTGTCGGGCGCCCTTCAACGCGGCAAGGAGATCGCGCTCGATCAT 1003
Db 1193 GAGCTGAGGACACATGACTGGAGTTCTCAGCGCGCAATGAGAAAGCTGTGAAATGT 1252
QY 1004 TCATCGCGGACGATCGGGTTTCTGGACATGCGCGGCGGTGGTGA 1049
Db 1253 TYATTGATGAAAGATAAGCTATTGATATATTTCAAGGTTGTGA 1298

RESULT 6
US-09-449-335-1
; Sequence 1, Application US/09449335
; Patent No. 6303365
; GENERAL INFORMATION:
; APPLICANT: Busch, Marco
; APPLICANT: Hain, Rudiger
; APPLICANT: Martin, William
; APPLICANT: Tietjen, Klaus
; APPLICANT: Kloti, Andreas
; TITLE OF INVENTION: Method of determining the activity of
; TITLE OF INVENTION: 1-deoxy-D-xylulose-5-phosphate reductoisomerase and
; TITLE OF INVENTION: 1-deoxy-D-xylulose-5-phosphate synthase
; FILE REFERENCE: 2020US
; CURRENT APPLICATION NUMBER: US/09/449,335
; CURRENT FILING DATE: 1999-11-24
; EARLIER APPLICATION NUMBER: DE 199 35 967.9
; EARLIER FILING DATE: 1999-07-30
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1732
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (18)..(1448)
; OTHER INFORMATION: DNA encoding 1-deoxy-D-xylulose-5-phosphate
; OTHER INFORMATION: reductoisomerase
US-09-449-335-1

Query Match 10.5%; Score 123.6; DB 4; Length 1732;
Best Local Similarity 48.7%; Pred. No. 7.6e-15;
Matches 373; Conservative 0; Mismatches 384; Indels 9; Gaps 1;

QY 293 CGGCCATCTGCGGGCGCGCTGTCGCCGAATGCGGGCGCTGAAGCAGCGCGCA 352
Db 550 CCGGAATAGTAGTTGTGCGGGACTAAGCCTACGGTTGCTGCAATTTGAAGCAGGAAGG 609
QY 353 CGTGGCGCTCGCAACAAGGAAGCCTCTGACGGCAGGGCAACTCTCTGATCGGACGG 412
Db 610 ACATTGCTCTTGCAACAAGAGACATTAATCGCAGGTGTCTTTCGTCTTCCGCTG 669
QY 413 CCCAGGAGAACGGCCACGATCTCGCGGTGGACAGCAGCACTCGCGGTCTTTCAGG 472
Db 670 CCAACAACATAATGTAAGATTTCTTCGCGAGATTCAGAACATTTGCAATATTTCACT 729
QY 473 CGCTGGCGGCGGAGCACAGCGCTGCTGAGCGCTCATCATCAGCGCTCGCGGGG 532
Db 730 GTATTCAAGTTTGGCTTGAAGGGCTCTCGCGAAGATAATCTTTCATCTGCTGCTGGAG 789
QY 533 CGTTCGCGACTGGAGCCTCGAGCGCATCGCGCTCGACCCCTGCGGCGGCGCAGGCC 592
Db 790 CTTTGGGATTTGGCTGTCGAAGAAGCTAAGGAAGTTAAAGTAGCGGATCGCTTGAAGC 849
QY 593 ATCCCAACTGGTCCATGGCGGCGGATCTTCATCGACAGCGCTCGATGTTCAACAAGG 652
Db 850 ATCCCAACTGGAACATGGGAAGAAAATCACTGTGACCTCTGTAGCTTTTCAACAAGG 909
QY 653 CGCTCGAGCTGATCGAGACCGGGAATTTCTCGGCTTCGAGCGCGGACCGGATCGAGCGG 712
Db 910 GTCTTGAAGTCAATGAAGCGCATTTTGTGGAGCTGAGTATGAGATATAGATTTG 969

QY 713 TCGTCATCCGCAATCCATCGTCCATGCGATGGTGGCTTCTGCGACGGGGCCCTGATGG 772
Db 970 TCATTTCATCCGCAAGATATCATACATTCATCCATGATGAAACACAGGATTCATCTGTGCTTG 1029
QY 773 CCCATCTCGGCGCCGCGACATGCGCCACGCGCATCGGATTCGGCTGAACTGCGCGGGTC 832
Db 1030 CTCATTTGGTGGCTGATATGCGTTTACCGATTCTCTACACCATGTGATGCGCCGATA 1089
QY 833 GCGGC-----GAGGTGCGCGTGGCGCGGATCGACCTCGCACAGATTGCGAGCCTCA 883
Db 1090 GAGTTCCTTGTCTGAAGTAACTTGGCCAAAGACTTGACCTTTGCAAACTCGGTTCAATGA 1149
QY 884 CTTTCAGAAAGCTGACGAGGACGCTTTTCGGCCCTTGAGGCTTGGCGAGACGTCATG 943
Db 1150 CTTTCAAGAAACAGACAATGTGAATACCCATCCATGATCTTCTTATGCTGCTGGAC 1209
QY 944 CGGCGCGCGCTGTGCGGCGCGCTTCAACGCGGCGCAAGGAGATCGCGCTCGATCAT 1003
Db 1210 GAGCTGGAGCAATGACTGGAGTTCTCAGCGCGCGCAATGAGAAAGCTGTTGAAATGT 1269
QY 1004 TCATCGCGGACGATCGGCTTCTGACATGCGCGCGCTGGTTCGA 1049
Db 1270 TCATTGATGAAGATAAGCTATTGGATATCTTCAAGTTGTGA 1315

RESULT 7

US-09-480-921B-9
; Sequence 9, Application US/09480921B
; Patent No. 6387637
; GENERAL INFORMATION:
; APPLICANT: Levin, Joshua Z.
; APPLICANT: Budziszewski, Gregory J.
; APPLICANT: Potter, Sharon L.
; APPLICANT: Wegrich, Lynette M.
; TITLE OF INVENTION: Herbicide Target Genes and Methods
; FILE REFERENCE: PB/5-30780A
; CURRENT FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1434
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1434)
US-09-480-921B-9

Query Match 10.3%; Score 122; DB 4; Length 1434;
Best Local Similarity 48.6%; Pred. No. 1.5e-14;
Matches 372; Conservative 0; Mismatches 385; Indels 9; Gaps 1;
QY 293 CGGCCATCGTGGCGCGCGGCTCGTCCCGGAATGCGGCGCTGAAGACGCGCGCA 352
Db 533 CCGGAATAGTAGTTGTGCGGCACTGAAGCCTACGGTTGCTGCAATTGAAGCAGGAAGG 592
QY 353 CGCTGGCTCGCCAAAGGAAGCCTCGTGAAGCGGCAAGCACTCTGATCGGACGG 412
Db 593 ACATTGCTCTGCAAAACAAGACATAATATCGCAGGTGGCTTTCGTTCCGCTTG 652
QY 413 CCAGGAGAACGCGCGCAGATCCTCGCGTGGACAGGAGCACTCCGCGGCTTTTCAGG 472
Db 653 CCAACAACATATGTAAGATTCTTCCGCGAGATTCAACAACATTCGCCATATTTCACT 712
QY 473 CGCTGGCGGCGAGGACAGCGCTCGCTGAGCGCTCATCATCGGCGCTCGGCGGC 532
Db 713 GTATTCAAGTTTGGCTGGAAGGCGCTCTCGCAAGATAATCTTGACTGCTGTTGGAG 772
QY 533 CGTTCCGCGACCTCGAGCGCATCGGCGCTGACCGTGGCGGAGGCGGCGCC 592
Db 773 CTTTAGGATTTGGCTGTGCAAAAGCTAAAGAGTAAAGTAGCGGATGCGTTGAAGC 832

QY 593 ATCCCAACTGGTCCATGGGCCAGCGATCTCCATCGACAGCGCTCGATGTTTCAACAAG 652
Db 833 ATCCAACTGGAACATGGGAAAGAAATCACTGTGGACTCTCTAGCTTTTCAACAAG 892
QY 653 CGCTGAGCTGATGAGAGCGCGGAATTTCTGGCTTCAGCGGACCGGATCGAGCGG 712
Db 893 GTCTTGAAGTCAATGAAGCGCATTTATTTGTTGGAGCTGAGTATGACGATAGAGATTG 952
QY 713 TCGTCCATCCGCAATCCATCGCTCCATGCGATGTTGGCTTCTGCGACGGGCGCTGATGG 772
Db 953 TCATTTCATCTCAAGTATCATACATTCATGATGAAACACAGGATTCATCTGCTTG 1012
QY 773 CCATCTCGGCGCGCGGACATGCGCCACGCGCATCGGATTCGGCTGAACTGCGCGGTC 832
Db 1013 CTCATTTGGTGGCTGATATGCGTTTACCGATTCTCTACACCATGTCTATGCGCGGATA 1072
QY 833 GCG-----GCGAGGTGCGCGCGCGGATCGACCTCGCACAGATTGCGAGCCTCA 883
Db 1073 GAGTTCCTTGTCTGAAGTAACTTGGCCTAGACTTTCGACCTTTGCAAACTCGGTTCAATGA 1132
QY 884 CTTTCAGAAAGCTGACGAGGAAGCCTTTTCGGCGCTGAGGCTTGGCGAGAGCTCATGG 943
Db 1133 CTTTCAAGAAACAGACAATGTGAATACCCATCCATGATCTTCTTATGCTGCTGGAC 1192
QY 944 CGGCGCGCGCTGTGCGGCGCGCTTCAACGCGGCGCAAGGAGATCGCGCTCGATCAT 1003
Db 1193 GAGCTGGAGCAATGACTGGAGTTCTCAGCGCGCGCAATGAGAAAGCTGTTGAAATGT 1252
QY 1004 TCATCGCGGACGATCGGCTTCTGAGATGCGCGGCTGAGATGCGCGGCTGGTTCGA 1049
Db 1253 TTATTGATGAAGATAAGCTATTGGATATCTTCAAGGTTGTGA 1298

RESULT 8

US-09-491-362-1
; Sequence 1, Application US/09491362
; Patent No. 6281017
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lange, Bernd M
; TITLE OF INVENTION: 1-DEOXY-D-XYLULOSE-5-PHOSPHATE REDUCTOISOMERASE, AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: WSUR14977
; CURRENT APPLICATION NUMBER: US/09/491,362
; CURRENT FILING DATE: 2000-01-26
; EARLIER APPLICATION NUMBER: 60/118,349
; EARLIER FILING DATE: 1999-02-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1759
; TYPE: DNA
; ORGANISM: Mentha piperita
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (72)..(1496)
US-09-491-362-1

Query Match 8.4%; Score 99.2; DB 4; Length 1759;
Best Local Similarity 48.3%; Pred. No. 2.3e-10;
Matches 377; Conservative 0; Mismatches 388; Indels 15; Gaps 3;
QY 293 CGGCCATCGTGGCGCGCGGCTCGTCCCGGAATGCGGCGCTGAAGACGCGCGCA 352
Db 598 CGGAATTCGCGCTCGCAGGTTTGAAGCGGACAGTGGCTGCCATAGAGCTGGAAG 657
QY 353 CGCTGGCTCGCCAAAGGAAGCCTCGTGAAGCGGCAACTCTGATCGGACGG 412
Db 658 ACATTGCTTTGGCCATAAAGACACATAATCGCTGAGGCGCTTTGCTTCTCTT 717
QY 413 CCAGGAGAACGCGCGGCGGCTCGCTGCGGTGAGACGAGCACTCGCGGCTTTTCAGG 472

Db 718 CAAGAAGCACAAAGTCAAGATTCTTCTCGCAGACTCCGAACATCTGCTATATTCAGT 777
Qy 473 CGCTGGCGGAGACACGCGCTCGCTCGAGCGGCTCATCATCAGCGCTCGCGCGGC 532
Db 778 GTATCAAGGCTTGCCAGAGTGTCTTGAGCGGTATATTTGACTGCATCGGAGAG 837
Qy 533 CGTTCGCGAGTGGAGCTCGAGCGCATCCGCGCTCGACCGTGGCGGAGCGAGGCC 592
Db 838 CTTTCAGGGATTGCGCGCTTGAGAAATTGAAAGAGTGAAGTAGCAGATGCTTTAAAGC 897
Qy 593 ATCCCACTGTCTCATGGCGCAGCGGATCTCCATCGACGAGCGCTCGATGTTTC---AACA 649
Db 898 ATTCCACTGGAATATGGGAAAAGAAATACAGTCCGACTTCTGCAACTCTTCTTTAACA 957
Qy 650 AGGCGCTCGAGTGCATCGAGACGCGCGAAATCTTCGCGCTTCGAGCGGACCGGATCGAGG 709
Db 958 AGGCGCTCGAAGTCATAAAGCTCACTATTGTTTGGGCGAGAATATGATGATTTGAGA 1017
Qy 710 CGGTGCTCCATCCGAATCCATGTCATGCGATGCGATGGTGGGCTTCTGCGACGGGGCGCTGA 769
Db 1018 TTGTTATTCAATTCGCCATCCATCACTACGTGCGTGGTCCGAGACACAGGATTCATCGGTGC 1077
Qy 770 TGGCCCATCTCGGCGCGCGCGACATCGCCAGCGCATCGGATTCGCGTGAAGTGGCGG 829
Db 1078 TAGCTCAATTAGATGGCGCGCGATGCGTTTGGCTTATCTGTACACCTTATCATGCGCAG 1137
Qy 830 -----GTGCGGCGGAGGTGCGCGCGCGCTTCGCGCTTCGAGCGGACCGGATTCGAGGCC 880
Db 1138 AGAGAGTCTACTGCTCGGAGATTACATGCGCTCGACTCGACCTCTGCAAGTGC---ATT 1194
Qy 881 TCACCTTCAGAGCTGACGAGAACGCTTTCGCGCGCTGAGGCTGCGGCGAGACGTCA 940
Db 1195 TACCATTCAAGAAGCGCGATTAACCGTGAATACCGCTATGCGATCTAGCCTATGCTGCTT 1254
Qy 941 TGGCGCGCGCGCTGCGCGCGCGCTTCGAGCGGCGGCAAGAGATCGCGCTCGATC 1000
Db 1255 GGAAGCGCGGAGCACCATGACCGAGTCTGAGCGCAGCTAATGAGAAACAGTCGAAA 1314
Qy 1001 ATTTATCGCGGAGCATCGGTTTCTGGACATGCGCGCGGTGTCGAGGAGACGCTCG 1060
Db 1315 TGTTCATCGAGAGAAATCGCTACCTCGACATTTTCAAGGCTGCTGAGCTTACATGCG 1374

RESULT 9

US-09-874-562-1

; Sequence 1, Application US/09874562

; Patent No. 6420159

; GENERAL INFORMATION:

; APPLICANT: Croteau, Rodney B

; APPLICANT: Lange, Bernd M

; TITLE OF INVENTION: 1-DEOXY-D-XYLULOSE-5-PHOSPHATE REDUCTOISOMERASE, AND

; FILE REFERENCE: WSUR17549

; CURRENT APPLICATION NUMBER: US/09/874,562

; PRIOR FILING DATE: 2001-06-04

; PRIOR FILING DATE: 2000-01-26

; PRIOR FILING DATE: 1999-02-03

; NUMBER OF SEQ ID NOS: 13

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 1759

; TYPE: DNA

; ORGANISM: Mentha piperita

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (72)..(1496)

US-09-874-562-1

Query Match

Best Local Similarity 48.3%; Pred. No. 2.3e-10;

Matches 377; Conservative 0; Mismatches 388; Indels 15; Gaps 3;

Qy 293 CGGCATCTGCGGCGCGCGGCTGTCGCCGAAATCGCGGCGCTGAAGCACGCGCGCA 352
Db 598 CGGAATTTGCTGGCTGTGAGGTTTGAAGCCGACAGTGGCTGCCATAGAACGCTGAAAGG 657
Qy 353 CGTGGCGCTGCCCAACAAAGAAAGCTCTGTCAGCGGAGGCAACTCCTGATGGGAGCG 412
Db 658 ACATTGCTTTGGCCAAATAAGAGACACTAATACGCTGGAGGCGCTTTTGTCTCTCTCTTG 717
Qy 413 CCCAGGAGAGCGGCGCCACGATCTGCGCGGTGGACAGCGGAGCAGCTCCGCGGTCTTTCAGG 472
Db 718 CAAGAAGCACAAAGTCAAGATTCTTCTCGACAGCTCCGAACATCTGCTATATTTCACT 777
Qy 473 CGCTGGCGGCGGAGACACGCGCTGCGTTCGAGCGGCTCATCATCAGCGGCTCGCGCGGC 532
Db 778 GTATCAAGGCTTCCGAGAGTGTCTTTGAGCGGTATATTTTACTGCTATGAGGAG 837
Qy 533 CGTTCGCGAGTGGAGCTCGAGCGCATCCGCGCTCGACCGTGGCGGAGCGGAGGCC 592
Db 838 CTTTCAGGGATTGCGCGCTTGAGAAATTGAAAGAGTGAAGTAGCAGATGCTTTAAAGC 897
Qy 593 ATCCCACTGTCTCATGGCGCAGCGGATTCCTCATCGACGCGCTCGATGTTTC---AACA 649
Db 898 ATTCCACTGGAATATGGGAAAAGAAATACAGTCCGACTTCTGCAACTCTTCTTTAACA 957
Qy 650 AGGCGCTCGAGTGCATCGAGACGCGCGAAATCTTCGCGCTTCGAGCGGACCGGATTCGAGG 709
Db 958 AGGCGCTCGAAGTCATAAAGCTCACTATTGTTTGGGCGAGAATATGATGATTTGAGA 1017
Qy 710 CGGTGCTCCATCCGAATCCATGTCATGCGATGGTGGGCTTCTGCGACGGGGCGCTGA 769
Db 1018 TTGTTATTCAATTCGCCATCCATCACTACGTGCGTGGTGGCTTATCTGTACACCTTATCATGCGCAG 1077
Qy 770 TGGCGCGCGCGCTGCGCGCGCGCTTCGAGCGGCGGCAAGAGATCGCGCTGAAGTGGCGG 829
Db 1078 TAGCTCAATTAGATGGCGCGCGATGCGTTTGGCTTATCTGTACACCTTATCATGCGCAG 1137
Qy 830 -----GTGCGGCGGAGGTGCGCGCGCGCTTCGCGCTTCGAGCGGACCGGATTCGAGGCC 880
Db 1138 AGAGAGTCTACTGCTCGGAGATTACATGCGCTCGACTCGACCTCTGCAAGTGC---ATT 1194
Qy 881 TCACCTTCAGAGCTGACGAGAACGCTTTCGCGCGCTGAGGCTGCGGCGAGACGTCA 940
Db 1195 TACCATTCAAGAAGCGCGATTAACCGTGAATACCGCTATGCGATCTAGCCTATGCTGCTT 1254
Qy 941 TGGCGCGCGCGCTGTCGCGGCGCGCTTCGAGCGGCGGCAAGAGATCGCGCTCGATC 1000
Db 1255 GGAAGCGCGGAGCACCATGACCGAGTCTGAGCGCAGCTAATGAGAAACAGTCGAAA 1314
Qy 1001 ATTTATCGCGGAGCATCGGTTTCTGGACATGCGCGCGGTGTCGAGGAGACGCTCG 1060
Db 1315 TGTTCATCGAGAGAAATCGCTACCTCGACATTTTCAAGGCTGCTGAGCTTACATGCG 1374

RESULT 10

US-07-945-283-1/c

; Sequence 1, Application US/07945283

; Patent No. 5352596

; GENERAL INFORMATION:

; APPLICANT: Cheung, Andrew K.

; APPLICANT: Wesley, Ronald D.

; TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants

; TITLE OF INVENTION: Involving The Ep0 and LIT Genes

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Curtis P. Ribando

; STREET: 1815 No. 5352596th University Street

; CITY: Peoria

; STATE: IL

; COUNTRY: USA

; ZIP: 61604

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/945,283
FILING DATE: 19920911

CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Ribando, Curtis P
REGISTRATION NUMBER: 27976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 309-685-4011 ext.513
TELEFAX: 309-685-4128

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 8438 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Pseudorabies virus

FEATURE:

NAME/KEY: CDS

LOCATION: 622..6495

FEATURE:

NAME/KEY: variation

LOCATION: replace(1099, "g")

FEATURE:

NAME/KEY: variation

LOCATION: replace(1267, "t")

FEATURE:

NAME/KEY: variation

LOCATION: replace(1381, "c")

FEATURE:

NAME/KEY: variation

LOCATION: replace(1566, "c")

FEATURE:

NAME/KEY: variation

LOCATION: replace(7010, "g")

US-07-945-283-1

Query Match 6.3%; Score 74.2; DB 1; Length 8438;
Best Local Similarity 45.1%; Pred. No. 9,4e-06;
Matches 320; Conservative 0; Mismatches 383; Indels 6; Gaps 1;

QY 10 CTGTGATCTTTTGGGCGACCGGCTCCATCGGCGAATCCACCTTCGACCTCGTCATGCGG 69

Db 4977 CTGCGGCACATCGGCGACGCCATGGCGCCAGGACCGCTCTGGGCGCTGCCGACGCG 4918

QY 70 AAGGCGGCGCGAGCGTTCCGACCGCTGCTGACCGGGGGGCGCAACATCGGCGCA 129

Db 4917 GTGAGCGCGGTGCCATGAGCGCGCGCTACGATCGCACCCAGAACCTTCATCTGCAG 4858

QY 130 CTGGCCGAATGTCGCGTGAAGCGGAGCTTGCGCTCACCGCGCATGAGGACTGC 189

Db 4857 AGCCTCCGCGGCGCTACGCGACATGGCTACCGGGCGCGCGGACCCCGCGCC 4798

QY 190 CTGGCGGCGCTCGCGAGCGCTGGCGGAGCGGACCGAGTCCGCGGGGGGCGGAG 249

Db 4797 GGGAGGCGCGCTGCGAGGCGCTTGCGCCCGCGCTCCGCGCGCGCTTGCGCCCGCGGAG 4738

QY 250 GCCATCGGCGGCG 309

Db 4737 CCGCGCGCGTCCCGCGAGCTGGCGGACGCTTGCCTGCGCTGCGCGCGCGCTC 4678

QY 310 GCGGGCGCTGCTCGCGGAAATGGGGCGGTGAAGCACGCGCGCGCGCGCTGCCAAC 369

Db 4677 GAGCGCGTCTGCGCTGCCGCTCCGCGTCCGCGCGCGCGCGCGCGCGCGCGCTC 4618

QY 370 AAGGAAAGCTGTGACGCGCAGGCGCACTCTGTATGCGGACGCGCGCGCGCGCGCC 429

Db 4617 GGGCGCGCGCTGCTCGAGGAGGTGACCGCGCGCGCTGCTCGCGTCCGACGCGATCCCC 4558

QY 430 ACGATCCTCGCGGTGACAGCGAGGACTCCGCGGTCTTTTTCAGGCGCTGCGGCGGAGGAC 489

Db 4557 GGGCGCGCGCGCGCGAGCGGAGGCGCGGACTCTGTGCGCTCTGCGCGCGCGCC 4498

QY 490 ACGGCTGCTGCGAGCGGCTCATCATCAGCGGCTCCGCGGCGCGGCTTCCGCGACTGGAGC 549

Db 4497 GTGGCGCGCGCTGCTGCGCTACAGCGTGGACGGGCGCGCGCGCGCGCGCGCGCGCG 4438

QY 550 CTCGAGCGCATCCGCGCGCTGACCGTGGCGCGGAGGCGGAGCGGCGGCGGCGGCGGCG 609

Db 4437 TACGCGCGCGCGCTCTTTCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGCGCGCGCG 4378

QY 610 GCGCAGCGGATCTCCATCGACAGCGCTCGATGTTCAACAAGCGCGCTCGAGCTGATCGAG 669

Db 4377 GCG 4324

QY 670 ACGCGCGAATCTTCGCGCTTCGAGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGG 718

Db 4323 CAGCGCGCGCTGCTGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4275

RESULT 11

US-09-000-016-1

; Sequence 1, Application US/090000016

; Patent No. 6143541

; GENERAL INFORMATION:

; APPLICANT: AKIRA ARISAWA et al.

; TITLE OF INVENTION: GENE ENCODING A PROTEIN HAVING ASYMMETRIC

; TITLE OF INVENTION: HYDROLASE ACTIVITY FOR 4-SUBSTITUTED 1,4-DIHYDROPYRIDINE DE

; TITLE OF INVENTION: ITS EXPRESSION PRODUCT

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.

; STREET: 2033 K Street, N.W., #800

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20006

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: Wordperfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/000,016

; FILING DATE: January 30, 1998

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Warren M. Cheek, Jr.

; REGISTRATION NUMBER: 33,367

; REFERENCE/DOCKET NUMBER:

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-721-8200

; TELEFAX: 202-721-8250

; TELEX:

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2809 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: genomic DNA

; ORIGINAL SOURCE:

; ORGANISM: Streptomyces viridosporus

; STRAIN: A-914

; ORGANISM: Streptomyces antibioticus


```
; STRAIN:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 338...2539
; IDENTIFICATION METHOD: E
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2540...2809
; IDENTIFICATION METHOD: P
; US-09-000-016-1
;
; Query Match 6.1%; Score 72.2; DB 3; Length 2809;
; Best Local Similarity 43.1%; Pred. No. 2.1e-05;
; Matches 347; Conservative 0; Mismatches 458; Indels 0; Gaps 0;
;
; QY 267 CGACCGCGCGCGGCTGATGCGGACCTGTCGGCCATGTCGGCGCGCGCGGCTCGTGCCCGG 326
; DB 1786 CGACGAGGTTTCGGCGCGGATCCAGGCCGACAGGCGCTCCAGCAGACCTGATCGCCGA 1845
;
; QY 327 AATCGCGGCGCTGAAGCAGCGCGCGCGGCTGCGGCTGCGCAAGAAAGCTCGTGAC 386
; DB 1846 CCGGCTCGGTCGAGTTCGGCGCTCCAGCAGTGGCGCGCACACGAGCGGCTCAC 1905
;
; QY 387 GGCAGGGCAACTCCTGATGCGGAGCGGCCAGGAGAGCGGCCACCATCTCGCGGTGA 446
; DB 1906 CAGCAGCTGACCTACCGCAACCTCGGCACCCAGGAGCTCAGCTGAAGCTGACGTGAC 1965
;
; QY 447 CAGCGAGCACTCCGCGGTCTTTCAGGCGCTGCGGCGGCGAGGACACGCGCTGCGTGAGCG 506
; DB 1966 CGCCACCGACCCAAAGGCAAGCGCGCGCGCGGCTTCTTTCAGCGTGGCGCCACCAC 2025
;
; QY 507 GGTATCATCAGCGGCTCGCGGCGGCGGTTCGCGGAGCTGAGGCTCGAGCGCATCGCGCG 566
; DB 2026 GGTGACGCTCGCGGCGGCGGCGGCGGCTCGCTGACATGACCGCGGACACCGCGGCTCG 2085
;
; QY 567 CTGACCGTGGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 626
; DB 2086 CGGCACGCTGAGCGGCGGCTGCTGCGGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGG 2145
;
; QY 627 CGACGCGGCTGATGTTTCAACAGGCGGCTGAGCTGATCGACGCGCGGCGGCGGCGGCGG 686
; DB 2146 CCGCAGCGCGCGCGGCTGCGAGCGGAGTTCGAGTTCGAGCGTACCGCTCGCGGCGCAT 2205
;
; QY 687 CTTCGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 746
; DB 2206 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 2265
;
; QY 747 GGGCTTCTGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 806
; DB 2266 GGGCTCGGCGCGGCTTACGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 2325
;
; QY 807 CGGATTGCGGCTGAACTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 866
; DB 2326 GGGCACCTACCTGTTGAGCTCTCTGATGCGCAAGGACTTCGGGAGCTCAAGGCGGCGAT 2385
;
; QY 867 ACAGATTGCGGCTGACCTTCCAGAGCGCTGACGAGGAGCGCTTTTCGCGGCGGCGGCGG 926
; DB 2386 CGACTGGCTGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2445
;
; QY 927 TGCGCGAGAGGCTATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 986
; DB 2446 ACGCACACCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2505
;
; QY 987 GATCGGCGCTCGATCATTTTATCGCGGAGCGATCTGCGGTTCTTCTGAGCATCGCGGCGG 1046
; DB 2506 GACCATCGGCTACACCTACGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2565
;
; QY 1047 CGAGGAGAGCGCTCGCGGCGGCTTTCG 1071
; DB 2566 CCGCGGCGGCTTCCCGGGGACCTTCG 2590
```

RESULT 12

```
US-09-514-340-1
; Sequence 1, Application US/09514340
; Patent No. 6361987
; GENERAL INFORMATION:
; APPLICANT: Akira ARISAWA et al.
; TITLE OF INVENTION: GENE ENCODING A PROTEIN HAVING ASYMMETRIC
; HYDROLASE ACTIVITY FOR 4-SUBSTITUTED 1,4-DIHYDROPYRIDINE
; ITS EXPRESSION PRODUCT
;
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
; STREET: 2033 K Street, N.W., #800
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/514,340
; FILING DATE: 28-Feb-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/000,016
; FILING DATE: January 30, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee Cheng
; REGISTRATION NUMBER: 40,949
; REFERENCE/DOCKET NUMBER: <Unknown>
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-721-8200
; TELEFAX: 202-721-8250
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2809 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Streptomyces antibioticus
; STRAIN: <Unknown>
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 338...2539
; IDENTIFICATION METHOD: E
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2540...2809
; IDENTIFICATION METHOD: P
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
; US-09-514-340-1
;
; Query Match 6.1%; Score 72.2; DB 4; Length 2809;
; Best Local Similarity 43.1%; Pred. No. 2.1e-05;
; Matches 347; Conservative 0; Mismatches 458; Indels 0; Gaps 0;
;
; QY 267 CGACCGCGCGCGGCTGATGCGGACCTGTCGGCCATGTCGGCGCGCGGCGGCTCGTGCCCGG 326
; DB 1786 CGACGAGGTTTCGGCGCGGATCCAGGCCGACAGGCGCTCCAGCAGACCTGATCGCCGA 1845
;
; QY 327 AATCGCGGCGCTGAAGCAGCGCGCGGCGGCTGCGGCTGCGCAAGAAAGCTCGTGAC 386
; DB 1846 CCGGCTCGGTCGAGTTCGGCGCTCCAGCAGTGGCGCGCACACGAGCGGCGGCTCAC 1905
;
; QY 387 GGCAGGGCAACTCCTGATGCGGAGCGGCCAGGAGAGCGGCCACCATCTCGCGGTGA 446
; DB 1906 CAGCAGCTGACCTACCGCAACCTCGGCACCCAGGAGCTCAGCTGAAGCTGACGTGAC 1965
```


QY 447 CAGCAGCACTCCGGGCTTTTACGGCGCTGGCGGGGAGGACACAGCGGCTTCGTCGAGCG 506
Db 1966 CGCCACCAGCCCAAGGCAAGGGCGCCCGCGGCTTCTTACGCTGGCGCGCACACCAC 2025
QY 507 CGTCATCATCAGCGCTCCGGCGGCGCTTCCGGGACTGGAGCCTTCGAGCGCATCCGCGC 566
Db 2026 GGTACCTCCCGGCGGCGGCGGCTTCCGTCGATGACGCGCCGACACCGGCTCGG 2085
QY 567 CTGACCGCTCCGAGCGGCGGCGGCTTCCCAACTGGTTCATGGCCACGCGGATCTCCAT 626
Db 2086 CGGCACGCTGACGCGGCTACTCGGCTAGCTGGTCCGACGCGGCGGCGGCGACAGCT 2145
QY 627 CGACAGCCCTCGATGTTTCAACAGGCGCTCGAGCTGATGAGAGCGCGCGAATTTTCGG 686
Db 2146 CCGACGCGCGCGGCTGACGCGGAGTGGAGTCTGACGCTGACCGTCCGCGCACAT 2205
QY 687 CTTGAGCGCGGACGATCGAGGCGGCTGCTCATCCGCAATCCATCGTCCATGCGATGCT 746
Db 2206 CGCGCGGACGCGCAAGCCACGACCGCAACACTCACCAGCTGATCGGCTACGCGGCGCT 2265
QY 747 GGGCTTCTGCGACGCGGCGCTGATGGCCCATCTCGGCGCGCGGCGGACATGCGCCACGCGCAT 806
Db 2266 GGGCTCGCGCGGCTTACGCGCGCGCGCGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2325
QY 807 CGGATTCGCGCTGAACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 866
Db 2326 GGGCACTTACCTGCTGGTACTCTGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2385
QY 867 ACAGATTCGAGGCTCAGCTTCCAGAGGCTGACGAGGAAAGCTTTCGCGGCGGCTGAGGCT 926
Db 2386 CGACTGGCTGCTCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2445
QY 927 TCGCGGAGACCTGATGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2445
Db 2446 ACGCACCAACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2505
QY 987 GATCGCGCTGATCATTTTCATCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1046
Db 2506 GACCATCGCTACACTACGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2565
QY 1047 CGAGGAGAGCTCGCGGCGGCTTCG 1071
Db 2566 CGCGCGGCTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2590

RESULT 13
US-08-242-677-1
; Sequence 1, Application US/08242677
; Patent No. 5677143
; GENERAL INFORMATION:
; APPLICANT: Gaynor, Richard B
; APPLICANT: Wu, Foon W.
; TITLE OF INVENTION: Cellular Nucleic Acid Binding Protein
; TITLE OF INVENTION: and Uses Thereof in regulating Gene Expression and in the
; TITLE OF INVENTION: Treatment of AIDS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/242,677
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

NAME: Mayfield, Denise L.
REGISTRATION NUMBER: 33,732
REFERENCE/DOCKET NUMBER: UTSD:401
TELEPHONE: 713-787-1400
TELEFAX: 713-789-2679
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5173 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..4863
US-08-242-677-1

Query Match 6.1%; Score 71.6; DB 1; Length 5173;
Best Local Similarity 48.9%; Pred. No. 2.8e-05;
Matches 222; Conservative 0; Mismatches 229; Indels 3; Gaps 1;

QY 158 CGAGCTTCCGCTACCCGCGATGAGGACTGCTCCCGCGCTGCGGAGGCGCTGGCGG 217
Db 341 CGAGCTGGCGGCGGCTGGCTGAGAGGCGCTGCTGCGCGGCTGCTGCGCGGCGGCGG 400
QY 218 GGACGGGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 277
Db 401 CGCTGGCGGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 460
QY 278 CC---GACTGGAACCATGTCGGCCATCTGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 334
Db 461 GCGAGGAGCGGCGGCTACTGAGAGCGGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 520
QY 335 CGCTGAACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 394
Db 521 GGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 580
QY 395 AACTCTGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 454
Db 581 GACTCTCCAGTGTGTCCTCAATGTCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 640
QY 455 ACTCGCGGCTTTTCAGGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 514
Db 641 TGGCGCGGCTGGGCGGCTCCCTGGGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 700
QY 515 TCACGCGCTCGGCGGCGGCGGCTTCGCGGACTGGAGCTCGAGCGCATCCGCGGCTGCACCG 574
Db 701 GCGCGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 760
QY 575 TGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 608
Db 761 AGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 794

RESULT 14
US-07-642-734C-3
; Sequence 3, Application US/07642734C
; Patent No. 5824513
; GENERAL INFORMATION:
; APPLICANT: Katz, L
; APPLICANT: Donadio, S
; APPLICANT: Mcalpine, J B
; TITLE OF INVENTION: Recombinant DNA Method for Producing
; TITLE OF INVENTION: Erythromycin Analogs
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edward H. Gorman
; STREET: Abbott Laboratories D377/AP6D-2 One Abbott
; CITY: Abbott Park
; STATE: IL
; COUNTRY: US

```

1 ZIP: 60064-3500
2
3 COMPUTER READABLE FORM:
4 MEDIUM TYPE: Floppy disk
5 COMPUTER: IBM PC compatible
6 OPERATING SYSTEM: PC-DOS/MS-DOS
7 SOFTWARE: PatentIn Release #1.0, Version #1.25
8 CURRENT APPLICATION DATA:
9 APPLICATION NUMBER: US/07/642,734C
10 FILING DATE: 17-JAN-91
11 CLASSIFICATION: 435
12 ATTORNEY/AGENT INFORMATION:
13 NAME: Danckert, Andreas M
14 REGISTRATION NUMBER: 32652
15 REFERENCE/DOCKET NUMBER: 4952.US.01
16 TELECOMMUNICATION INFORMATION:
17 TELEPHONE: 708-937-9396
18 TELEFAX: 708-938-2623
19 INFORMATION FOR SEQ ID NO: 3:
20 SEQUENCE CHARACTERISTICS:
21 LENGTH: 20235 base pairs
22 TYPE: nucleic acid
23 STRANDEDNESS: double
24 TOPOLOGY: unknown
25 MOLECULE TYPE: DNA (genomic)
26 HYPOTHETICAL: NO
27 ANTI-SENSE: NO
28 ORIGINAL SOURCE:
29 ORGANISM: Saccharopolyspora erythraea
30 STRAIN: NRRL 238
31 FEATURE:
32
33 NAME/KEY: CDS
34 LOCATION: 19..10722
35 OTHER INFORMATION: /codon_start= 19
36 OTHER INFORMATION: /function= "gene eryA"
37 OTHER INFORMATION: /product= "eryA ORF2 encoding modules 3 & 4 for
38 OTHER INFORMATION: 6-deoxyerythronolide B"
39 FEATURE:
40
41 NAME/KEY: misc_feature
42 LOCATION: 19..4470
43 OTHER INFORMATION: /function= "approximate span of
44 OTHER INFORMATION: module 3"
45 FEATURE:
46
47 NAME/KEY: misc_feature
48 LOCATION: 97..1482
49 OTHER INFORMATION: /function= "approximate span of
50 OTHER INFORMATION: beta-ketoacyl ACP synthase of module 3"
51 FEATURE:
52
53 NAME/KEY: misc_feature
54 LOCATION: 1693..2670
55 OTHER INFORMATION: /function= "approximate span of
56 OTHER INFORMATION: acyltransferase domain module 3"
57 FEATURE:
58
59 NAME/KEY: misc_feature
60 LOCATION: 3406..3921
61 OTHER INFORMATION: /function= "approximate span of
62 OTHER INFORMATION: beta-ketoreductase domain of module 3"
63 FEATURE:
64
65 NAME/KEY: misc_feature
66 LOCATION: 4171..4428
67 OTHER INFORMATION: /function= "approximate span of
68 OTHER INFORMATION: acyl carrier domain of module 3"
69 FEATURE:
70
71 NAME/KEY: misc_feature
72 LOCATION: 4471..10722
73 OTHER INFORMATION: /function= "approximate span of
74 OTHER INFORMATION: module 4"
75 FEATURE:
76
77 NAME/KEY: misc_feature
78 LOCATION: 4471..5847
79 OTHER INFORMATION: /function= "approximate span of
80 OTHER INFORMATION: beta-ketoacylACPSynthase domain of module"
81 FEATURE:
82
83 NAME/KEY: misc_feature

```

```

; OTHER INFORMATION: /function- "approximate span of
; OTHER INFORMATION: acyl carrier domain of module 6"
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 19492..20235
; OTHER INFORMATION: /function- "approximate span of
; OTHER INFORMATION: thioesterase domain of module 6"
US-07-642,734C-3

```

[illegible]

RESULT 15

```

RES001 13
US-08-439-009A-3
; Sequence 3, Application US/08439009A
; Patent No. 6004787
; GENERAL INFORMATION:
; APPLICANT: Donadio, S
; APPLICANT: Katz, L
; APPLICANT: Mcapine, J B
; TITLE OF INVENTION: Method of Directing Biosynthesis of
; TITLE OF INVENTION: Specific Polyketides
; NUMBER OF SEQUENCES: 27

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Steven F. Weinstein
STREET: Abbott Laboratories D377/AP6D-2 One Abbott
STREET: Park Rd
CITY: Abbott Park
STATE: IL
COUNTRY: US
ZIP: 60064-3500

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/439,009A
FILING DATE: 11-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Casuto, Dianne
REGISTRATION NUMBER: 40,943
REFERENCE/DOCKET NUMBER: 4952.US.D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847-938-3137
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 20235 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Saccharopolyspora erythraea
STRAIN: NRRL 238

FEATURE:
NAME/KEY: CDS
LOCATION: 19..10722 /codon_start= 19
OTHER INFORMATION: /function= "gene erya"
OTHER INFORMATION: /product= "eryA ORF2 encoding modules 3"
OTHER INFORMATION: 6-deoxyerythronolide B"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19..4470
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 97..1482
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoacyl ACP synthase of module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1693..2670
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyltransferase domain module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3406..3921
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoreductase domain of module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4171..4428
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyl carrier domain of module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4471..10722
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: module 4"
FEATURE:

```

; NAME/KEY: misc_feature
; LOCATION: 4471..5847
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: beta-ketoacylACPSynthase domain of module"
; FEATURE:
;
; NAME/KEY: misc_feature
; LOCATION: 6054..7026
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyltransferase domain of module 4"
; FEATURE:
;
; NAME/KEY: misc_feature
; LOCATION: 7165..9216
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: dehydratase and enoylreductase domains m"
; FEATURE:
;
; NAME/KEY: misc_feature
; LOCATION: 9433..9984
; OTHER INFORMATION: /function= "approximate span
; OTHER INFORMATION: beta-ketoreductase of module 4"
; FEATURE:
;
; NAME/KEY: misc_feature
; LOCATION: 10225..10483
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyl carrier domain of module 4"
; FEATURE:
;
; NAME/KEY: CDS
; LOCATION: 10723..20235
; OTHER INFORMATION: /codon_start= 10723
; OTHER INFORMATION: /function= "gene-eryA"
; OTHER INFORMATION: /product= "orf3 encoding modules 5 & 6
; OTHER INFORMATION: 6-deoxyerythronolide B formatio"
; FEATURE:
;
; NAME/KEY: misc_feature
; LOCATION: 10723..15165
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: module 5"
; FEATURE:
;
; NAME/KEY: misc_feature
; LOCATION: 10831..12174
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: beta-ketoacylACPSynthase domain of modul"
; FEATURE:
;
; NAME/KEY: misc_feature
; LOCATION: 12379..13350
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyltransferase domain of module 5"
; FEATURE:
;
; NAME/KEY: misc_feature
; LOCATION: 14062..14610
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: beta-ketoreductase of module 5"
; FEATURE:
;
; NAME/KEY: misc_feature
; LOCATION: 14857..15114
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyl carrier domain of module 5"
; FEATURE:
;
; NAME/KEY: misc_feature
; LOCATION: 15166..20235
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: module 6"
; FEATURE:
;
; NAME/KEY: misc_feature
; LOCATION: 15172..16569
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: beta-ketoacylACPSynthase domain of modul"
; FEATURE:
;
; NAME/KEY: misc_feature
; LOCATION: 16768..17721
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyltransferase domain of module 6"
; FEATURE:
;
; NAME/KEY: misc_feature
```

```

; LOCATION: 18379..18921
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: beta-ketoreductase domain of module 6"
; FEATURE:
;
; NAME/KEY: misc_feature
; LOCATION: 19149..19398
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyl carrier domain of module 6"
; FEATURE:
;
; NAME/KEY: misc_feature
; LOCATION: 19492..20235
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: thioesterase domain of module 6"
; FEATURE:
;
; US-08-439-009A-3

Query Match      5.9%; Score 69.4; DB 3; Length 20235;
Best Local Similarity 43.7%; Pred. No 7.4e-05;
Matches 304; Conservative 0; Mismatches 391; Indels 0; Gaps 0;

QY 381 CGTGACGGCAGGCAACTCTGTATGCGGAGCGGCCCGCCAGGAGAACGGCCACGATCCTGCC 440
      |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 5829 CGTGGTTCGGGGCCCGCCGCGCTCGCCGAGACCTCGCGGACACACCCCGAGCTCGG 5888

QY 441 GGTGGACAGGACACTCGCGGTCTTTAGGCGCTGGCGGGGAGAGACAGCGCTGCGT 500
      || || || || || || || || || || || || || || || || || || || ||
Db 5889 CTTGACCGACGCGCGGTGGACGCTCGCGACCGCGAGGGCGGTTGCGACGTGGACCGCC 5948

QY 501 CGAGCGCGTCATCATCACGCGCTCCGCGGGCGCTTCCGGAGCTGGAGCCCTCGAGGCGCAT 560
      || || || || || || || || || || || || || || || || || || || ||
Db 5949 CGTGCTCGGGGACGACCGCGGGCGTGTGCGGAGCTGGAGCGGCTGCCCGAGGGCGG 6008

QY 561 CCGCGCTCGACCGTGGCGGAGCGCGAGGCCCATCCCAACTGTGTCTCATGGCGGACGCGAT 620
      || || || || || || || || || || || || || || || || || || || ||
Db 6009 CCGGTCGGCGGCGCGCTCGCGCGGTGACTCCGCGCGCGCAAGCCGCTCTGTGTTT 6068

QY 621 CTCCTATCGACGCGCTCGATGTTCAACAAGGCGCTCGAGCTGATCGAGACGCGCGAATT 680
      || || || || || || || || || || || || || || || || || || || ||
Db 6069 CCCCAGCGAGGGCGCGAGTGGGTGCGCATGGCACCGCATCTGCTGGAATCTCTCCGAGGT 6128

QY 681 CTTCCGCTTCGAGCGGACCGGATCGAGGCGGTCTGCTCCATCCGCAATCCATCTCCATGC 740
      || || || || || || || || || || || || || || || || || || || ||
Db 6129 GTTCGCGGAGTCGATGAGCGGTGCGCGGAGCGCTCTCGCGCGACACCGACTTGAAGTT 6188

QY 741 GATGGTGGGCTTCGCGAGGGGGCGCTGATGCGCCCATCTCGGCGCGCGCGCGCGCGCGCA 800
      || || || || || || || || || || || || || || || || || || || ||
Db 6189 GCTCGACGTCGTCGCGGGGACGCGGTCGCCGACCGCGACGAGCGGCTCGACGTGTCCA 6248

QY 801 CGCCATCGGATTCGCGCTGNACTGGCGGGGTCCGCGGAGGTCCCGCGGATCGA 860
      || || || || || || || || || || || || || || || || || || || ||
Db 6249 GCCGGTGCTTCTTCGATCATGTTCTCGTGGCGGAGCTGTGGCGCGCGCGCGCGGTGAC 6308

QY 861 CCTCGCACAGATTGCGAGCGCTCACTTCCAGAAAGCCTGACGAGGAACGCTTTCGGGCGCT 920
      || || || || || || || || || || || || || || || || || || || ||
Db 6309 CCGCGCGCGCTGTCGCGCCACTCGAGGGCGAGATCGCGCGCGCGCGCGCGCGCGCG 6368

QY 921 GAGGCTTGGCGGAGAGCTGATGGCGGGCGCGGCTGTGCGGGCGCGCGCTTCAACCGGGC 980
      || || || || || || || || || || || || || || || || || || || ||
Db 6369 GCTGTCGTGGAAGCGCGCGGAGGTGTGGTGGCCCTCGCGCAGCGCAGGTGTGCGCGAGCT 6428

QY 981 CRAAGGAGATCGCGCTCGATCATTTTCATCGCGCGGACGATCGGGTTTCTGGACATGGCGG 1040
      || || || || || || || || || || || || || || || || || || || ||
Db 6429 CGACGACCGAGGGCGGATGTTGTCGGTGGCGGCGTCCCGCGCGAGCTGGAGACCGTGCT 6488

QY 1041 GGTGGTCGAGGAGACGCTCGCGGGCGCTTTCGACCG 1075
      || || || || || || || || || || || || || || || || || || || ||
Db 6489 CGCGCGCTGGAGCGGCGGTCTCGCGGTGGCGCGCG 6523
```

Search completed: April 22, 2003, 04:38:44
Job time : 4312 secs

STIC-Biotech/ChemLib

CRFF

9/4/25

4/15 01+02

From: Saidha, Tekchand
Sent: Sunday, April 13, 2003 10:11 AM
To: STIC-Biotech/ChemLib
Subject: 09/673198 - sequence search request

09/673198

Please search the data base and interference files for SEQ ID NOS : 30 and 31.

Thank you,

Jekchand Saidha

Primary Examiner

Art Unit 1652, CM1, Room No. 10D05

Mail Box 10D01

(703) 305-6595

Point of Contact:
Toby Port
Technical Info. Specialist
CM1 6A04
703-305-3534

30-394 aa
31-1182 na

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 4/15
Date Completed: 4/22
Searcher Prep/Review: 12
Clerical: _____
Online time: 12

TYPE OF SEARCH:
NA Sequences: 1
AA Sequences: 1
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: es
WWW/Internet: _____
Other (specify): _____

